





[illegible][illegible]







[illegible][illegible]

[illegible]

07	1812	ccaaagcgaaactatccaaattgtttgaaatattctatgagcgcgagatccaaagac	18711
Db	2427	ccaaagccgaactatccaaattgtttgaaatattctatgagcgcgagatccaaagac	24866
07	1972	gacgggaactccaaactccaaacaaagagatgagagagagagagagagagagag	19311
Db	2487	gacitggaaattccaaactccaaacaaagagatgagagagagagagagagagagag	25466
07	1932	ctccctgctgagagagagatggaattgacattgcgaagatccctcccgatgagctg	19411
Db	2547	cttccttgcctgacacgaacgattggaattgacattgcgcgaactccctcccgatgac	26066
07	1942	gcttgatgaagagatataagcttcgaagctccaaagagagagagagagagagagag	20511
Db	2607	gcttgatgaagagatataagcttcgaagctccaaagagagagagagagagagagag	26666
07	2052	cgatttgaagcttccgagagagagagagagagagagagagagagagagagagag	21111
Db	2667	cgatttgaagcttccgagagagagagagagagagagagagagagagagagagag	27266
07	2112	agcaaacagagatccaaagagagagagagagagagagagagagagagagagagag	21711
Db	2727	agcaaacagagatccaaagagagagagagagagagagagagagagagagagagag	27866
07	2172	tctaacgagatccaaag	22311
Db	2787	tctaacgagatccaaag	28466
07	2232	ctggcgagagatccaaagagagagagagagagagagagagagagagagagagag	22911
Db	2847	ctggcgagagatccaaagagagagagagagagagagagagagagagagagagag	29066
07	2292	gaaactcttccaaagatctgaaactccaaacccctccctcgaacttgagagagagag	23511
Db	2907	gaaactcttccaaagatctgaaactccaaacccctccctcgaacttgagagagagag	29666
07	2352	atccggcgagagatccaaagagagagagagagagagagagagagagagagagagag	24111
Db	2967	atccggcgagagatccaaagagagagagagagagagagagagagagagagagagag	30266
07	2412	ttcgattccgaagagatctgaaactccaaacccctccctcgaacttgagagagagag	24711
Db	3027	ttcgattccgaagagatctgaaactccaaacccctccctcgaacttgagagagagag	30866
07	2472	gcaagcatttcgagagagatctgaaactccaaacccctccctcgaacttgagagagag	25311
Db	3087	gcaagcatttcgagagagatctgaaactccaaacccctccctcgaacttgagagagag	31466
07	2532	caaacgcgagatccaaag	25911
Db	3147	caaacgcgagatccaaag	32066
07	2592	ctcatttcgagagagatccaaagagagagagagagagagagagagagagagagagag	26511
Db	3207	ctcatttcgagagagatccaaagagagagagagagagagagagagagagagagagag	32666
07	2652	gatttcgag	27111
Db	3267	gatttcgag	33266
07	2712	ctgag	27711
Db	3327	ctgag	33866
07	2772	gatttcgag	28311
Db	3387	gatttcgag	34466
07	2832	ttcgattccgaagagatctgaaactccaaacccctccctcgaacttgagagagagag	28911
Db	3447	ttcgattccgaagagatctgaaactccaaacccctccctcgaacttgagagagagag	35066





[illegible][illegible]











[illegible][illegible]

[illegible][illegible]

```
|||||
1b 721 GCTGAATGAG AGTGGATGCTGTACAGTACATGGCTTGGAGGCGGTTC 780
QY 2492 CCAAGCAAAAGTGTGTATTCCTTACCAAGCTTGTATCCGCTTAGCATTTCA 2451
|||||
1b 781 CGAAGCAAAAGTGTGTATTCCTTACCAAGCTTGTATCCGCTTAGCATTTCA 840
QY 2452 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2511
|||||
1b 841 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 900
QY 2512 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2571
|||||
1b 901 AAGATATCTTCTTATGAGAAATAGAGTGTGTGAGTAAATGATTCGCAAAATTC 960
QY 2572 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2631
|||||
1b 961 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1020
QY 2632 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2691
|||||
1b 1021 TGGTCAATACCTGCTTCAATGCTGAGACCGGAGAGTGTGTGAGTAAATGATTC 1080
QY 2692 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2751
|||||
1b 1081 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1140
QY 2752 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2811
|||||
1b 1141 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1200
QY 2812 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2871
|||||
1b 1201 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1260
QY 2872 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2931
|||||
1b 1261 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1320
QY 2932 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 2991
|||||
1b 1321 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1380
QY 2992 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 3051
|||||
1b 1381 CCAAGCTTCTCCCTGGAGAGTAAATGATTCGCAAAATTCATCTGATCACTAGC 1440
```

Search completed: August 16, 2002 07:58:11  
Job time: 13595 sec













[illegible]

Db	2701	caaaagctgtctcgggaacacactccgcgcagctacgagaaacccggggcccaagttccca	2760
Gy	2761	gagtcgggcaaaatgctctgaatctggtggtgaatgagctgcttccaaagagctccg	2820
Gy	2821	agcgcctctatgctgctgaacgtctctcaacacacctctggcttgcggctccagagctcc	2880
Db	2821	agcgcctctatgctgctgaacgtctctcaacacacctctggcttgcggctccagagctcc	2880
Gy	2881	aagagctgtct	2940
Db	2941	gccctcagatcgggtccaaagatgctgctgaagatctctgctgctgctgctgctgctgca	3000
Gy	2941	gccctcagatcgggtccaaagatgctgctgaagatctctgctgctgctgctgctgctgca	3000
Db	2941	gccctcagatcgggtccaaagatgctgctgaagatctctgctgctgctgctgctgctgca	3000
Gy	3001	tag 3003	
Db	3001	tag 3003	
RESULT 3			
AAZ58405			
AAZ58405	standard; cDNA; 3618 BP.		
AAZ58405			
23-MAY-2000	(first entry)		
RNA encoding fumonisin esterase- <i>ami</i> amino polyol amine oxidase fusion.			
XX			
XX	Amino polyol amine oxidase: APO; fumonisin esterase: mycotoxin;		
XX	transgenic plant; detoxification, animal feed, silage;		
XX	glutathione S-transferase: alpha-amylase, signal peptide; maize;		
XX	SS.		
XX	Chimeric - Schistosoma japonicum.		
OS	Chimeric - Exophiala spinifera		
XX			
FT	Key	location/Qualifiers	
FT	sig_peptide	1..687	
FT		/tag= a	
FT	mat_peptide	/product= "GST _ polylinker"	
FT		677..3615	
FT		/tag= b	
FT	misc_feature	/product= "fumonisin-esterase-K-triAPO fusion"	
FT		2227..2229	
FT		/tag= c	
FT		/note= "extra lysine codon"	
XX	WO200004159-A1.		
PN			
XX	27-JAN-2000.		
XX			
PE	08-JUL-1999;	99WO-0515454.	
XX			
PR	15-JUL-1998;	98US-0092936.	
XX	21-MAY-1999;	99US-0145391.	
XX			
PA	(PIONE) PIONEER HI-PPED INT INC.		
XX	(CUPA) CUPAGEN CORP.		
XX			
PI	Duvick JP, Gilliam JT, Maddox JR;		
XX			
DR	WPI; 2000-182425/16.		
XX			
DP	P-PDR; AAV58416.		
XX			
PT	New isolated polynucleotides, polypeptides useful for detecting and		
PT	degrading fumonisin or structurally related mycotoxin in processed		
PT	grain or in silage -		

XX  
1-5 Example 13; Page 115-120; 154pp; English

The present sequence is that of a polynucleotide encoding a fusion protein (see AAF56946) composed of glutathione S transferase and Escherichia spinifera liponisin esterase (EspI) joined via a peptide linker to K. LAWA. (see AAF52907), a truncated, but still functional, amino polyol amino oxidase of E. spinifera. The construct was designed for expression in Escherichia coli of a hybrid protein having EspI and APO activity. The invention describes APO polynucleotides (see AAF26383-87) and polypeptides (see AAF56800-05). E. spinifera and Rhinodielis acrocytus. The polynucleotides are used to transform plant cells normally susceptible to Fusarium or other toxin producing fungus infection. Also provided are methods for expressing APO, optionally as a fusion protein with liponisin esterase, in transgenic plants, prokaryotic and non-plant eukaryotic systems. Methods for detection of grain, grain processing, silage, food crops and in animal feed and rumen microorganisms are also disclosed.

Sequence 1018 BP, 984 A, 625 C, 677 A, 679 C, 681 A, 683 A, 685 A, 687 A, 689 A, 691 A, 693 A, 695 A, 697 A, 699 A, 701 A, 703 A, 705 A, 707 A, 709 A, 711 A, 713 A, 715 A, 717 A, 719 A, 721 A, 723 A, 725 A, 727 A, 729 A, 731 A, 733 A, 735 A, 737 A, 739 A, 741 A, 743 A, 745 A, 747 A, 749 A, 751 A, 753 A, 755 A, 757 A, 759 A, 761 A, 763 A, 765 A, 767 A, 769 A, 771 A, 773 A, 775 A, 777 A, 779 A, 781 A, 783 A, 785 A, 787 A, 789 A, 791 A, 793 A, 795 A, 797 A, 799 A, 801 A, 803 A, 805 A, 807 A, 809 A, 811 A, 813 A, 815 A, 817 A, 819 A, 821 A, 823 A, 825 A, 827 A, 829 A, 831 A, 833 A, 835 A, 837 A, 839 A, 841 A, 843 A, 845 A, 847 A, 849 A, 851 A, 853 A, 855 A, 857 A, 859 A, 861 A, 863 A, 865 A, 867 A, 869 A, 871 A, 873 A, 875 A, 877 A, 879 A, 881 A, 883 A, 885 A, 887 A, 889 A, 891 A, 893 A, 895 A, 897 A, 899 A, 901 A, 903 A, 905 A, 907 A, 909 A, 911 A, 913 A, 915 A, 917 A, 919 A, 921 A, 923 A, 925 A, 927 A, 929 A, 931 A, 933 A, 935 A, 937 A, 939 A, 941 A, 943 A, 945 A, 947 A, 949 A, 951 A, 953 A, 955 A, 957 A, 959 A, 961 A, 963 A, 965 A, 967 A, 969 A, 971 A, 973 A, 975 A, 977 A, 979 A, 981 A, 983 A, 985 A, 987 A, 989 A, 991 A, 993 A, 995 A, 997 A, 999 A, 1001 A, 1003 A, 1005 A, 1007 A, 1009 A, 1011 A, 1013 A, 1015 A, 1017 A, 1019 A, 1021 A, 1023 A, 1025 A, 1027 A, 1029 A, 1031 A, 1033 A, 1035 A, 1037 A, 1039 A, 1041 A, 1043 A, 1045 A, 1047 A, 1049 A, 1051 A, 1053 A, 1055 A, 1057 A, 1059 A, 1061 A, 1063 A, 1065 A, 1067 A, 1069 A, 1071 A, 1073 A, 1075 A, 1077 A, 1079 A, 1081 A, 1083 A, 1085 A, 1087 A, 1089 A, 1091 A, 1093 A, 1095 A, 1097 A, 1099 A, 1101 A, 1103 A, 1105 A, 1107 A, 1109 A, 1111 A, 1113 A, 1115 A, 1117 A, 1119 A, 1121 A, 1123 A, 1125 A, 1127 A, 1129 A, 1131 A, 1133 A, 1135 A, 1137 A, 1139 A, 1141 A, 1143 A, 1145 A, 1147 A, 1149 A, 1151 A, 1153 A, 1155 A, 1157 A, 1159 A, 1161 A, 1163 A, 1165 A, 1167 A, 1169 A, 1171 A, 1173 A, 1175 A, 1177 A, 1179 A, 1181 A, 1183 A, 1185 A, 1187 A, 1189 A, 1191 A, 1193 A, 1195 A, 1197 A, 1199 A, 1201 A, 1203 A, 1205 A, 1207 A, 1209 A, 1211 A, 1213 A, 1215 A, 1217 A, 1219 A, 1221 A, 1223 A, 1225 A, 1227 A, 1229 A, 1231 A, 1233 A, 1235 A, 1237 A, 1239 A, 1241 A, 1243 A, 1245 A, 1247 A, 1249 A, 1251 A, 1253 A, 1255 A, 1257 A, 1259 A, 1261 A, 1263 A, 1265 A, 1267 A, 1269 A, 1271 A, 1273 A, 1275 A, 1277 A, 1279 A, 1281 A, 1283 A, 1285 A, 1287 A, 1289 A, 1291 A, 1293 A, 1295 A, 1297 A, 1299 A, 1301 A, 1303 A, 1305 A, 1307 A, 1309 A, 1311 A, 1313 A, 1315 A, 1317 A, 1319 A, 1321 A, 1323 A, 1325 A, 1327 A, 1329 A, 1331 A, 1333 A, 1335 A, 1337 A, 1339 A, 1341 A, 1343 A, 1345 A, 1347 A, 1349 A, 1351 A, 1353 A, 1355 A, 1357 A, 1359 A, 1361 A, 1363 A, 1365 A, 1367 A, 1369 A, 1371 A, 1373 A, 1375 A, 1377 A, 1379 A, 1381 A, 1383 A, 1385 A, 1387 A, 1389 A, 1391 A, 1393 A, 1395 A, 1397 A, 1399 A, 1401 A, 1403 A, 1405 A, 1407 A, 1409 A, 1411 A, 1413 A, 1415 A, 1417 A, 1419 A, 1421 A, 1423 A, 1425 A, 1427 A, 1429 A, 1431 A, 1433 A, 1435 A, 1437 A, 1439 A, 1441 A, 1443 A, 1445 A, 1447 A, 1449 A, 1451 A, 1453 A, 1455 A, 1457 A, 1459 A, 1461 A, 1463 A, 1465 A, 1467 A, 1469 A, 1471 A, 1473 A, 1475 A, 1477 A, 1479 A, 1481 A, 1483 A, 1485 A, 1487 A, 1489 A, 1491 A, 1493 A, 1495 A, 1497 A, 1499 A, 1501 A, 1503 A, 1505 A, 1507 A, 1509 A, 1511 A, 1513 A, 1515 A, 1517 A, 1519 A, 1521 A, 1523 A, 1525 A, 1527 A, 1529 A, 1531 A, 1533 A, 1535 A, 1537 A, 1539 A, 1541 A, 1543 A, 1545 A, 1547 A, 1549 A, 1551 A, 1553 A, 1555 A, 1557 A, 1559 A, 1561 A, 1563 A, 1565 A, 1567 A, 1569 A, 1571 A, 1573 A, 1575 A, 1577 A, 1579 A, 1581 A, 1583 A, 1585 A, 1587 A, 1589 A, 1591 A, 1593 A, 1595 A, 1597 A, 1599 A, 1601 A, 1603 A, 1605 A, 1607 A, 1609 A, 1611 A, 1613 A, 1615 A, 1617 A, 1619 A, 1621 A, 1623 A, 1625 A, 1627 A, 1629 A, 1631 A, 1633 A, 1635 A, 1637 A, 1639 A, 1641 A, 1643 A, 1645 A, 1647 A, 1649 A, 1651 A, 1653 A, 1655 A, 1657 A, 1659 A, 1661 A, 1663 A, 1665 A, 1667 A, 1669 A, 1671 A, 1673 A, 1675 A, 1677 A, 1679 A, 1681 A, 1683 A, 1685 A, 1687 A, 1689 A, 1691 A, 1693 A, 1695 A, 1697 A, 1699 A, 1701 A, 1703 A, 1705 A, 1707 A, 1709 A, 1711 A, 1713 A, 1715 A, 1717 A, 1719 A, 1721 A, 1723 A, 1725 A, 1727 A, 1729 A, 1731 A, 1733 A, 1735 A, 1737 A, 1739 A, 1741 A, 1743 A, 1745 A, 1747 A, 1749 A, 1751 A, 1753 A, 1755 A, 1757 A, 1759 A, 1761 A, 1763 A, 1765 A, 1767 A, 1769 A, 1771 A, 1773 A, 1775 A, 1777 A, 1779 A, 1781 A, 1783 A, 1785 A, 1787 A, 1789 A, 1791 A, 1793 A, 1795 A, 1797 A, 1799 A, 1801 A, 1803 A, 1805 A, 1807 A, 1809 A, 1811 A, 1813 A, 1815 A, 1817 A, 1819 A, 1821 A, 1823 A, 1825 A, 1827 A, 1829 A, 1831 A, 1833 A, 1835 A, 1837 A, 1839 A, 1841 A, 1843 A, 1845 A, 1847 A, 1849 A, 1851 A, 1853 A, 1855 A, 1857 A, 1859 A, 1861 A, 1863 A, 1865 A, 1867 A, 1869 A, 1871 A, 1873 A, 1875 A, 1877 A, 1879 A, 1881 A, 1883 A, 1885 A,

July 1971, 97.68; Score 2932; DB 21; Length 3618;

Best and available: 100.00, Fed. No. 0,  
Not above 100.00 (average) in a  
district.

Matches: 0, Conservatives: 0, Mistatches: 0, Indels: 0, Gaps: 0,

07	72	caactactactcaagatctatgctctggaatggaatcagacacactactactgctccggg	131
10	667	gactctactactcaagatctatgctctggaatggaatcagacacactactactgctccggg	746
07	112	caactactactcaagatctatgctctggaatggaatcagacacactactactgctccggg	191
10	747	caactactactcaagatctatgctctggaatggaatcagacacactactactgctccggg	806
07	192	gagctactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	251
10	807	gagctactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	866
07	252	gagctactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	311
10	867	gagctactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	926
07	412	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	371
10	927	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	986
07	372	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	431
10	987	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	1041
07	482	atatactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	491
10	1047	atatactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	1101
07	492	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	551
10	1107	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	1161
07	552	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	611
10	1167	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	1221
07	612	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	671
10	1227	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	1281
07	672	gagctactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	731
10	1287	gagctactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	1341
07	742	caactactactactgctccggtgcttggtaagagctttgaaagcactgacatgctcc	791

[illegible]







[illegible][illegible]













[illegible][illegible]

	RESIDUE	L1
AZ58495		
ID	AAV58495 standard; cDNA; 1673 bp.	
XX		
AC	AAV58495;	
XX		
PE	7-8-MAY-2000 (first entry)	
XX		
NP	Active polyclonal anti oxalase (alpha mating factor signal peptide) RNA	
XX		
KM	Active polyclonal anti oxalase (NMC-1 immunisation agent) rabbit	
RW	transgenic plants detoxification; animal feed; silage;	
FJ	selectable marker; alpha mating factor; signal peptides ss-	
XX		
OS	Chimeric - Saccharomyces cerevisiae.	
CS	chimeric - Exophiala spiliotera.	
XX		
FM	Key Location/Qualifiers	
FT	CDS	1..1665
FT	/CDS	/taag-a
FT	sig_peptide	1..267
FI	/ftatg-b	
FT	/product "yeast alpha mating factor secretion	
FT	mat-peptide	signal"
ET		268..1662
ET	/taag-a	
FT	/product "K::TRAPAO"	
FT	misc_feature	268..270
FI	/ftatg-d	
FT	/note "extra lysine"	
XX		
FN	WO200004159-A1.	
PD	27-JAN-2000.	
XX		
PB	08-JUL-1999;	99MO-US15454.
XX		
PR	15 JUL-1998;	98DU-0092946.
XX	21-MAY-1999;	99SE-0135891.
PA	(FTECH) FINDER III BREED INT INC.,	
XX	(CURA) CURAGEN CORP.	
PI	Dawick JP, Gilliam JT, Madrox JR,	
XX		
DP	WP1: 2000-182425/16.	
DR	p-PSP95 AAV58910.	
EX		

We isolated polymeric fibrils, polypeptides useful for detecting and degrading tumourin or structurally related mycotoxin in processed grain or in silage .







10	AAZ58400 standard: cDNA; 2679 bp.
XX	AAZ58400:
XX	24-MAY-2000 (first entry)
XX	DNA encoding GST fused with amino polyol amino oxidase fusion.
XX	
XX	Amino polyol amino oxidase: APOA; human; myofibroblast; transgenic plant; detoxification; animal feed; sludge; selectable marker; glutathione S transferase; GST; ss.
XX	chimeric - Schistosoma japonicum.
XX	chimeric - Esophila spalliera.
XX	key: /**at = amino polyol amino oxidase
XX	single-peptide 1.487
XX	/**at =
XX	mat-peptide /**protein = "ast fusion + polylinker"
XX	6000.2076 /**at = b
XX	fusion = "K-trAAPAO"
XX	misc_feature 608..690
XX	/**at =
XX	feature "extra lysine"
XX	W200004159-A1.
XX	27-JAN-2000.
XX	08-JUL-1999: 9060-US15454.
XX	15-JUL-1998: 9805-US062836.
XX	21-MAY-1999: 9805-US05941.
XX	(PION-) P1 NTP H-18PDS INT 1N
XX	(CURA-) CURAEN COMP.
XX	David JP, Gilliam JT, Mathew JR.
XX	WPI: 2000-16245/16.
XX	P-PSDB: AAY58911.
XX	New isolated polynucleotides, polypeptides useful for detecting and
XX	detecting tumours or structurally related myofibroblast in processed
XX	urine or in sludge.
XX	Example 2 Page 60-62, figure 10a-10c.
XX	
XX	The present sequence is that of a polynucleotide encoding a
XX	GST-K-trAAPAO fusion protein composed of glutathione S transferase
XX	and K-trAAPAO, a transgene for human polyol amino oxidase (see also AAZ58400) of Esophila spalliera. The construct
XX	was designed for expression of K-trAAPAO in Esophila cells.
XX	K-trAAPAO is capable of degrading tumours and related myofibroblasts.
XX	The invention provides a GST-K-trAAPAO fusion protein, a polypeptide
XX	polypeptides (see AAY58908-95) of Esophila spalliera and K-trAAPAO
XX	derivatives. The polynucleotides are used to transcribe plasmids
XX	normally susceptible to fusion of other toxin-producing fungus
XX	infections. Also provided are methods for expression of AAZ58400
XX	transgenic plants, prokaryotic and non-plant eukaryotic systems,
XX	methods for detoxification of urine, urine processing, animal feed
XX	compos and in animal feed and tumor microorganisms are also disclosed.
XX	AZAO polynucleotide can also be used as a selectable marker.
XX	Sequence 2079 bp, 537 A, 490 T, 547 G, 595 T, 0 other.

Werner's Match	46.48	Score 1993	EB 215	Length 2079
Best Local Similarity	100.00	Prod No	0	
Matches 1993	0	Mismatches	0	Gaps 0

[illegible]



[illegible]

PA	(PION-) PIONEER HI-BRED INT. INC.
XA	(CUHA-) CORAGEN CORP.
ET	Danish JT, William JT, Madsen JP;
XX	
XX	Wt% 2000-197425/16.
NR	P-PNDR; AAV58909.
PT	New isolated polynucleotides, polyprotein genes for deterring and
ET	inhibiting infection of structurally related species in processed
PT	grain or in silage
XX	
PS	Example 7, page 81-83; 154pp; English.
XX	
CC	The present sequence is that of a polynucleotide encoded a truncated,
CC	but still functional, amino poly(1) amine oxidase (K1TRAPV), see
CC	AAV58909) of Euphorbia spinirostris. The polynucleotide was derived from
CC	clone KONG 345-5.5 (see AAV58909) to which has added a 5' lysine codon
CC	(nucleotides 1-3); Since many amine oxidases have a positively charged
CC	amino acid near the N-terminus and upstream of a dinucleotide binding
CC	site, The construct allowed heterologous expression of TRAPV in
CC	E. coli pasteurized and mated. The invention provides APAO polynucleotides
CC	(see AA58381-87) and polypeptides (see AAV58900-05) of E. spinifera and
CC	Rhinocallia strobilacea. The polynucleotides are used to transform
CC	plant cells normally susceptible to Fusarium or other toxin-producing
CC	fungal infections. Also provided are methods for expressing APAO in
CC	transgenic plants, prokaryotic and non-plant eukaryotic systems.
CC	Methods for detoxification of grain, grain processing, silage, food
CC	crops and in animal feed and rumen microorganisms are also disclosed.
CC	APAO polynucleotide can also be used as a selectable marker.
XX	
CC	Sequence: 1392 nt, 315 A, 370 C, 395 G, 292 T; 0 other:
Query Match                  46.4%; Score 1392; H# 21; Length 1392;	
Best local similarity    100.0%; Prod. No. 0;	
Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps	
OY	1612 aaagcaacgcttggaagactgatattatggcgttttgttagacgaattttagaacgaca 1671
DB	1 aaaagcaacgcttggaagactgatattatggcgttttgttagacgaattttagaacgaca 60
OY	1672 cgcagaactccagccgcgcgtctcttcctgcctcgcttttagaggcatatgatcttatgg 1731
DB	61 cgcagaactccagccgcgcgtctcttcctgcctcgcttttagaggcatatgatcttatgg 120
OY	1732 ggaagacctctgaqgcatacaatccaggtcccgagcaagaaactatctgaagacctctgg 1791
DB	121 ggaagacctctgaqgcatacaatccaggtcccgagcaagaaactatctgaagacctctgg 180
OY	1792 agctaatcaatgagagcaaaaaggctgattctccgatttctcagaaagtattatctt 1851
DB	161 agctaatcaatgagagcaaaaaggctgattctccgatttctcagaaagtattatctt 240
OY	1852 gaagagagctccagagaggaagctgtggaaattcaatccgaccagagcaagaaactatgac 1911
IL	241 gagagagagctccagagaggaagctgtggaaattcaatccgaccagagcaagaaactatgac 300
OY	1912 actaagactctctatggagatctctctctctctctctctctctctctctctctctctct 1971
DB	301 actaagactctctatggagatctctctctctctctctctctctctctctctctctctct 460
OY	1972 ctctctccgctgatctctcagatctatctcaaaagagagatctctcaagactccaagacgg 2031
DB	361 ctctctccgctgatctctcagatctatctcaaaagagagatctctcaagactccaagacgg 420
OY	2032 cctctcagagagagagctctgagaaattcttttttttttttttttttttttttttttt 2091
DB	421 cctctcagagagagagagctctgagaaattcttttttttttttttttttttttttttttt 480
OY	2092 ttgcctactggtctcagagctctgagagcaagctatcacagagctctctctctctctctct 2151
DB	481 ttgcctactggtctcagagctctgagagcaagctatcacagagctctctctctctctctct 540



Sequence version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

us nucleotide - nucleotide search using sw model

Run on: August 16, 2002 06:16:40 : Search time 6.99e seconds  
(without alignments)  
Iteration 10, 111111, 2011, 41444/5000

Hit(s): US-09-658-835c-24

Percent score: 100%  
1 alignment(s) with score 100% (100% of the sequence)

Scoring table:  
Gapop 60 0, Gapext 60 0

Searched: 484544 seqs, 122816752 residues

Word size: 3

Total number of hits satisfying chosen parameters: 767056

Minimum hit seq length: 0

Maximum hit seq length: 200

Post-processing: listing first 45 summaries

Database:

Issued Patents NA:  
1. US-09-352-159-24  
2. US-09-352-159-24  
3. US-09-352-159-24  
4. US-09-352-159-24  
5. US-09-352-159-24  
6. US-09-352-159-24  
7. US-09-352-159-24  
8. US-09-352-159-24  
9. US-09-352-159-24  
10. US-09-352-159-24  
11. US-09-352-159-24  
12. US-09-352-159-24  
13. US-09-352-159-24  
14. US-09-352-159-24  
15. US-09-352-159-24  
16. US-09-352-159-24  
17. US-09-352-159-24  
18. US-09-352-159-24  
19. US-09-352-159-24  
20. US-09-352-159-24  
21. US-09-352-159-24  
22. US-09-352-159-24  
23. US-09-352-159-24  
24. US-09-352-159-24  
25. US-09-352-159-24  
26. US-09-352-159-24  
27. US-09-352-159-24

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARY

Result No.	Score	Query Match	Length (bp)	HL	Description
1	2003	100.0	3003	4	US-09-352-159-24
2	8003	100.0	3003	4	US-09-352-159-24
3	2932	97.5	1618	4	US-09-352-159-24
4	2932	97.5	1618	4	US-09-352-159-24
5	1428	47.5	2976	4	US-09-352-159-24
6	1428	47.5	2976	4	US-09-352-159-24
7	1428	47.5	2976	4	US-09-352-159-24
8	1428	47.5	2976	4	US-09-352-159-24
9	1428	47.5	2976	4	US-09-352-159-24
10	1428	47.5	2976	4	US-09-352-159-24
11	1428	47.5	2976	4	US-09-352-159-24
12	1428	47.5	2976	4	US-09-352-159-24
13	1428	47.5	2976	4	US-09-352-159-24
14	1428	47.5	2976	4	US-09-352-159-24
15	1428	47.5	2976	4	US-09-352-159-24
16	1428	47.5	2976	4	US-09-352-159-24
17	1428	47.5	2976	4	US-09-352-159-24
18	1428	47.5	2976	4	US-09-352-159-24
19	1428	47.5	2976	4	US-09-352-159-24
20	1428	47.5	2976	4	US-09-352-159-24
21	1428	47.5	2976	4	US-09-352-159-24
22	1428	47.5	2976	4	US-09-352-159-24
23	1428	47.5	2976	4	US-09-352-159-24
24	1428	47.5	2976	4	US-09-352-159-24
25	1428	47.5	2976	4	US-09-352-159-24
26	1428	47.5	2976	4	US-09-352-159-24
27	1428	47.5	2976	4	US-09-352-159-24

#### ALIGNMENTS

```

RESULT 1
US-09-352-159-24
Sequence 24, Affiliation: 12, 00452159A
Patent No. 6211434
GENERAL INFORMATION:
APPLICANT: Davick, Jonathan P.
APPLICANT: Gilliam, Jacob T.
APPLICANT: Masada, Joyce P.
TITLE OF INVENTION: Amino Acid and Peptide Polymers and Methods of Use
FIELD OF INVENTION: Polymers, peptides, and amino acids
PRIORITY CLAIM: 1134
PARENT APPLICATION: US-09-352-159-24
EARLIER FILING DATE: 1998-07-12
EARLIER APPLICATION NUMBER: 60/092,936
EARLIER FILING DATE: 1998-07-25
EARLIER APPLICATION NUMBER: 60/135,391
EARLIER FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FASTA for Windows Version 3.0
SEQ ID NO 24
LENGTH: 3003
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Sequence is barley alpha amylase signal sequence;
Patent No. 6211434
OTHER INFORMATION: csp1 mat: an artificial spacer sequence and
OTHER INFORMATION: K:ITPAPO
FEATURE:
NAME/KEY: sig-peptide
LOCATION: (1)...(72)
OTHER INFORMATION: barley alpha amylase signal sequence
FEATURE:
NAME/KEY: misc-feature
LOCATION: (73)...(1575)
OTHER INFORMATION: csp1 mat
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1612)...(3000)
OTHER INFORMATION: K:ITPAPO
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(3000)
NAME/KEY: misc-feature
LOCATION: (1612)...(1614)
OTHER INFORMATION: Extra lysine

```











[illegible]

00	2547	ctcttctgtctgacgagagatctgacatctgactcttgcgaaatctctccgctctgattctctc	2605
07	1592	gctatctccaaatgacatagatccttcaataatctccagatgagagatctccagctcaatccgctcca	2051
06	2607	gctgctctgaagatgagatagctcttcataatctccaggctgagagatctccagctcaatccgctcca	2666
07	2052	cagctctgagctctcgcgacactatctgataatgagatcaaatctccctgctgctctccgagct	2111
06	2667	cagctctgagctctcgcgacactatctgataatgagatcaaatctccctgctgctctccgagct	2726
07	2112	agcaaacacatctcagcgcgctctgctgctctgctgctctgctgctctgctgctctgctgctctg	2171
06	2727	agcaaacacatctcagcgcgctctgctgctctgctgctctgctgctctgctgctctgctgctctg	2786
07	2172	ctctccacacatctcagcgcgctctgctgctctgctgctctgctgctctgctgctctgctgctctg	2231
06	2787	ctctccacacatctcagcgcgctctgctgctctgctgctctgctgctctgctgctctgctgctctg	2846
07	2232	ctctccacacatctcagcgcgctctgctgctctgctgctctgctgctctgctgctctgctgctctg	2291
06	2847	ctctccacacatctcagcgcgctctgctgctctgctgctctgctgctctgctgctctgctgctctg	2906
07	2292	ggaatctctctcagctctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	2351
06	2907	ggaatctctctcagctctcagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	2966
07	2352	atccgctgctatctgagctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct	2411
06	2967	atccgctgctatctgagctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct	3026
07	2412	ctcgtctctcagcact	2471
06	3027	ctcgtctctcagcact	3086
07	2472	gcaagcctctgagcgaatctctatctctgagctactatagagatagctctctctctctctctct	2531
06	3087	gcaagcctctgagcgaatctctatctctgagctactatagagatagctctctctctctctctct	3146
07	2532	cagagcctctctgagcgaatctctatctctgagctactatagagatagctctctctctctctctct	2591
06	3147	cagagcctctctgagcgaatctctatctctgagctactatagagatagctctctctctctctctct	3206
07	2592	ctcattctcagagataccagcactctcagcctcagctgagatctgctctctctctctctctctctct	2651
06	3207	ctcattctcagagataccagcactctcagcctcagctgagatctgctctctctctctctctctctct	3266
07	2652	ggtcgaagatcccgagcgaatctgctctcgaacgctcgaatgagatctgagcgaatctgct	2711
06	3267	ggtcgaagatcccgagcgaatctgctctcgaacgctcgaatgagatctgagcgaatctgct	3326
07	2712	ctgagcctctctcgcgcgagctctgctgagatcccgagcctcgcgcgagcctcgcgcgagcctcgc	2771
06	3327	ctgagcctctctcgcgcgagctctgctgagatcccgagcctcgcgcgagcctcgcgcgagcctcgc	3386
07	2772	ct	2831
06	3387	ct	3446
07	2832	ctgagcctctctcgcgcgagctctgctgagatcccgagcctcgcgcgagcctcgcgcgagcctcgc	2891
06	3447	ctgagcctctctcgcgcgagctctgctgagatcccgagcctcgcgcgagcctcgcgcgagcctcgc	3506
07	2892	cttgcgtctgagcctgagcgcctcttctctggaagctctctctggaagcctgagctctctctct	2951
06	3507	cttgcgtctgagcctgagcgcctcttctctggaagctctctctggaagcctgagctctctctct	3566
07	2952	gagctctgagcctgagcgcctcttctctggaagctctctctggaagcctgagctctctctct	3003
06	3567	gagctctgagcctgagcgcctcttctctggaagctctctctggaagcctgagctctctctct	3618





```

1 NAME/KEY: musc_1eat_170
2 LOCATION: (2773)
3 OTHER INFORMATION: KITAPAO
4 FEATURE:
5 NAME/KEY: CUS
6 LOCATION: (1)---(2773)
7 NAME/KEY: musc_1eat_170
8 LOCATION: (2773)---(1767)
9 OTHER INFORMATION: Extra lysine
10 US-09-652-159-25
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
263
```













```

Sequence ID Application US/2014/02197A
Patent No. 6211434
GENERAL INFORMATION:
APPLICANT: DUVICK, Jonathan P.
APPLICANT: Gilliam, Jacob L.
APPLICANT: Maddox, Joyce K.
TITLE OF INVENTION: Actin-Polyol Polymers
FIELD OF INVENTION: Polymericolides and Related Polypeptides and Methods of Use
FILE REFERENCE: 1124
CURRENT APPLICATION NUMBER: US/2014/02197A
EARLIER FILING DATE: 10-08-2012
EARLIER APPLICATION NUMBER: 62/092,936
EARLIER FILING DATE: 10-08-2012
EARLIER APPLICATION NUMBER: 60/135,391
EARLIER FILING DATE: 10-08-2012
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 16
LENGTH: 1673
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME: vvvv
LOCATION: (1)...(267)
OTHER INFORMATION: yeast alpha mating factor secreted signal
FEATURE:
NAME: vvvv
LOCATION: (1)...(1662)
US-00-000000-15-00-10

```

























was constructed by Dr. Claudio Schneider (LMB-Arca Science Park, Trieste, Italy).

BASE COUNT	120 g	110 g	150 g	73 g
BASELINE				

Best Local Similarity: 100.0%; Pred. No. 19;  
 Matches: 21; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

27 1261 tttagaatctgaagcttttc 1281  
 |||  
 34 427 tttagaatctgaagcttttc 407

Search completed: August 16, 2002, 08:56:27  
 CPU time: 10611 sec





401 TTAGTTTGAATATACGCTGAGCTGGCATATATGTAATTCATTC 450  
 151 HisLeuTyrAspGlyAlaSerPheAlaAlaAsnGlnAspValIleAlaVal 167  
 451 CAGCCTTACAGAGGCTTACCTTCCAGCCAAICAGAGATGTCATCCGCG 500  
 167 ThrIleAsnTyrArgThrAsnIleLeuGlyPheProAlaAlaProGln 184  
 501 GATTAATTAACATACAGAGCAACATTCGAGCTTCCTTCCTGCTCCGAG 550  
 184 GlnProIleThrGlnArgAsnLeuGlyPheLeuAspGlnIleArgPheAlaLeu 200  
 551 TTTCAALAACAAAGCAAAATCTGGCTTCCTACAGCAAAAGTTTGTTTG 600  
 201 AsnTyrValIleIleArgAsnIleAlaIlePheGlyGlyAspProArgVal 217  
 601 GATTCGGTACAGTCGAGACATCCGAGCTTCGAGCTGATCTCCCAAGG 650  
 217 ThrIlePheGlyGlnSerAlaGlyTyrAlaSerValAspValIleLeuLeu 234  
 651 CAAATAATTTGGGCAAGTGGGGGAGCAAGTGTGAGCTCTCTCTGA 700  
 234 HisSerMetProHisAsnProProPheArgAlaAlaIleGlnGluSerGly 250  
 701 GGTCTAATGCAACCAACCCCTCCGAGCAGCAATCATGGAGTCGGGT 750  
 251 ValAlaAsnTyrAsnProProGlySerAspGlnSerGlnProIlePheAsn 267  
 751 GTGGCTACATACAACTCCCAAGGCAATTTCTCCAACTTCGACAC 800  
 267 ThrValIleAlaIleLeuAsnGlyThrThrSerIleAspIleLeuSerGly 284  
 801 CAGCTGTCAGAGCTCTCAACTGACCAACCAATACATCATCTGAGTTGTA 850  
 284 CysArgArgValAspLeuAlaThrLeuMetAsnThrIleGlnGlnLeuGly 300  
 851 TGAATAACATTCGATCTGCTGATGAAACGATCCAGCAACTGGGA 900  
 301 LeuGlyPheGlnTyrThrLeuAspAsnValThrAlaValTyrArgSerGly 317  
 901 CTTCGGTTTGAGTACAGCTGGTGGCAACGTAAAGCGCTGTGACGGTTCTGA 950  
 317 GlnThrAlaArgIleThrGlyAspIleAlaArgValProValLeuValGly 334  
 951 AACAAATTCGAACAGGAGAAATGAGTGGGTAACCTGCTTCGGAGGA 1000  
 334 HisValAlaAsnAspGlyLeuLeuPheValLeuGlyGlnAsnAspThrGln 350  
 1001 GAGTGGCAAGCAGGAGCTCTCTTCCTGCTCGGGAGAAATGACACCCAA 1050  
 351 AlaTyrLeuGlnIleAlaIleProAsnGlnIleProAspIleuTyrGlnThrLe 367  
 1051 TATAATTCAGAGAGAGGTAAATCCCAATCAGCCGCACTTTACCAAGCTCT 1100  
 367 GluGlnValAlaTyrProIleGlySerProGlyIleGlySerProGlnAsp 384  
 1101 CATTGAGGCAATCCCATTCGATCCGAGGATCGCATCGCTCCCAAGATC 1150  
 384 IleIleAlaAlaIleGlnThrGlnValArgPheGlnGlyProSerAlaIle 400  
 1151 AGATTCGCGCTATTGAGACCGAGGTAAAGATTCGAGTGTCTTCGCGCAG 1200  
 401 ValAlaGlnAspSerArgAsnArgGlyIleProSerTrpArgTyrTyrGly 417  
 1201 GTGGCTTACAGCTTCGAGATTCGAGCTTTCTCTGGGTCGAGAGCTTAC 1250  
 417 HisAsnIleThrPheGlnAsnLeuGlnLeuPheProGlySerGlnValTyr 434  
 1251 CAAATGTAACCTTCGACAAATTCGAGCTTTCTCTGGGTCGAGAGCTTAC 1300  
 434 HisSerSerGlnValGlyMetValPheGlyThrTyrProValAlaSerAla 450  
 1401 ACAGCTGTCAATTCGAGATTCGCTTTCTTGGTAACTATCTGTGGCAACTGG 1350

451 ThrAlaLeuGlnAlaGlnThrSerTyrTyrMetGlnGlyAlaTrpAlaAl 467  
 1351 ACCGCTCTTGAGGCCAGCCAGCAAGCAAAATACATCCAGGCTGCTGGCG 1400  
 467 ArgPheAlaTyrAsnProMetAsnGlyProGlyTyrPheGlySerValProAsn 484  
 1401 CTTTGGCAAAACCCCTATTAAATGGAGCTGTGGTGGAAACAAATGTGGAA 1450  
 484 AlaAlaAlaLeuGlySerTyrGlyGlyAlaIleGlnValAspValSerPro 500  
 1451 TGGTACAGCTTGGTACCTAAAGCAAGCAAGCAAGCTTACGCTCTCTCCA 1500  
 501 AlaThrIleAspGlnArgGlyAlaLeuTyrThrArgTyrTyrThrGlnLe 517  
 1501 GCGAATAAGACCAAGCATGCTCTGTACACGCTTAAATATATGATCT 1550  
 517 GlnTyrThrIleAlaProArgPhePheGlyGlyGlySerGlyLeuIleSer 534  
 1551 GGGCAAAATGCGCGCGAGAAATTTGGGCAAGTAAAGCGGAGCAAGAG 1600  
 534 TyrGlyGlySerTyrAspAsnValAlaAspValValValValGlyValSer 550  
 1601 GCGAGAGCAACAAACCAACCTTCAAGCAAGCTGAGTGGTGGGCTGGC 1650  
 551 LeuSerGlyLeuGlnThrAlaArgValValGlnAlaIleGlyLeuSerGly 567  
 1651 TTGAGCGTTTGGACAGCGTACGTAAATTCACACCGCTGGTGTGTGTG 1700  
 567 SerValLeuGlnAlaMetAspArgValGlyGlyTyrThrLeuSerValG 584  
 1701 CTTCTCTTGGAGCGATGATGATGATGATGATGATGATGATGATGATG 1750  
 584 HisSerGlyProGlyTyrArgThrThrIleAsnAspGlyGlyValAlaTrp 600  
 1751 AATGAGCTGTGGTAAAGCAATTAACCAAGTGTGGTGTGGTGTGGTGT 1800  
 601 AsnAspSerAsnGlnSerGlnValSerArgLeuPheGlnArgPheHisLe 617  
 1801 AATGACAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1850  
 617 GlnGlnGlyGlyLeuGlnArgThrThrGlyAsnSerIleHisGlnAlaGln 634  
 1851 GAGAGCGCAAGCTTGAAGAGCGCAATGGAAATTCATCATCATCAATAC 1900  
 634 SerGlyThrThrThrThrAlaProTyrGlyAspSerLeuLeuSerGlnGln 650  
 1901 ACGGTACCAAGCTACAGTCTCTTAAGGTAAGCTCTTGTGAGTGGAGGAG 1950  
 651 ValAlaSerAlaLeuAlaGlnLeuLeuProValTyrProGlnLeuIleG 667  
 1951 GTTCAAGCTGACTTGGCAAGCTTCTGCTGATGGCTGAGCTGATGCA 2000  
 667 GlnHisSerLeuGlnAsnLeuAlaSerProGlnAlaTyrAsnArgLeu 684  
 2001 AGAGCATACCTTCAAGACCTCAAGAGCGCCCTAGGCGAAGGAGGCTGG 2050  
 684 SerSerValSerPheAlaHisTyrCysGlnGlyGlnLeuAsnLeuProAla 700  
 2051 ACAGTGTGAGCTTGGCGCAATAGTGTGAGCAAGCAAGCAAGCTTGGCT 2100  
 701 ValLeuGlyValAlaAsnGlnIleThrArgValAlaGlnGlyValGlnVal 717  
 2101 GTTCGCGCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2150  
 717 HisGlnIleSerMetLeuPheLeuThrAspTyrIleTyrSerAlaThr 734  
 2151 CCAAGCATCAAGCATGCTTTTCTTCAAGCAATATCAAGATGCGAGCTG 2200  
 734 TyrLeuSerAsnIlePheSerArgTyrTyrAspGlyGlyGlnTyrMetArg 750  
 2201 GTCTGAGTAATATTTCTTGGCAAGAAAGAAAGCGGAGGAGTATATGAGA 2250



1551	GGGCAATACGGCGGCGAGCAATTCGGGAAACAAAGAGCGAGAAAGCG	1600
534	LYSLGLYSERLYSASPASNVALIASPVALVALVALVALGLVALAGLY	550
1601	AGGAGAGACCAACACCAACACCTTGCGACGGTGTAGTGTGGCGCTGGC	1650
551	LEUSERGLYLEUCLUTHRALAATGLYSVALGLNALAALACLYLEUSERCY	567
1651	TTASGCGGTTTGAAAGAGAGAGAGCAAGAGAGAGAGAGAGAGAGAGAG	1700
567	SLGVALLEUGLVALIMELASPAVALGLYGLYSERLEUSERVALG	584
1701	CTCGCTTCTTGAGAGCGATGGATCGTGTACGGGAAAGAACTCGAGCTGAC	1750
584	INSERCTLYPROGLYARGTHRTHRILEASNASPLEUGLYALIAALATPRLIE	600
1751	AATCGAGTCTCCGCAAGAGTACTATACAGAGTCTAGAGAGAGAGAGAG	1800
601	ASPAASPSORASNOINSERGLYVALSORARGLEPHGLIARGPHGLISLE	617
1801	AATGACAGCAACCAAGAGCAAGATCCAGATTGTTGAAAGATTTCATT	1850
617	UGLUGLYLLEUGLILARGTHRTHRGLYASUSERTLEHISGLNALAGLA	634
1851	GGAGGGCGAGCTCCAGACGACAGATCGAATTCATATCATCATCAACCAAG	1900
634	SPGLIYTHRTHRTHRPRALAPROTYGTHYASPSERTLEHISERLEUGLIE	650
1901	AGGGGAAACCACTACAGCTCTCTATGCGACTCTCTGTAGAGCAAGAG	1950
651	VALIASERVALALEUAGLILEUENUPROVALTRPSEGLILEUILLGL	667
1951	GTTTCGAACTGCACTTGCGGCAACTCTCTCCCGCATGAGTCCAGCTATCGA	2000
667	UGLILHISERLEUGLILASPLEULYSASERTPROGLIALYLSARGLEUA	684
2001	AGAGCAATACCTTCAGAGCTCAAGCTCAAGCTCAAGCTCAAGCGAGCTCG	2050
684	SPSERVALSERPHEALIAHISTYCYSGULYSLILEASILEUPROALA	700
2051	ACAGTGTGAGTTCGGCGCACTACTGTGAAAGAAAGAACTAACTTGATCTGCT	2100
701	VALLEUGLYVALIASNOCTHETHARGALILEUENCTLYVALCTUAL	717
2101	GTTCTCTCCGCTAGCAAGCATCAACACCGCTCTGCTCTGGTGTGCAAGC	2150
717	AHLSGLIILIESEMERLEPHLEUETHASPYRTLEYSERLATHARG	734
2151	CGACGAGATCAGCATGCTTTTCTGCACGCACTATATCAAGAGTGTGACCG	2200
734	LYLEUSERASNILEPHESESPASFLYGLYSASPTLYGLIINTYRWEIARG	750
2201	GTCTCACTAATATTCTCTCGCAACAAAGAGCGCGGTACTIVAIAGCA	2250
751	CYSGLYTHRGLEMERLESTLEHISGLIIMELTSLYSGHILGCTVY	767
2251	TCTAAACATCTATCTACCTCATTTCTGATATCTCTGTAAGCAACTTCT	2300
767	IPROGLYSERVALHISLEASNTHRPROVALAAGLILEGLINSERLA	784
2301	TTCAGAGCTCAGTGCACCTTCACACGCGCTGCTGTAATTCAGACGTCGG	2350
784	IASERTGLYSGTHRVALARGSERLALASERGLYALVALIIPHARGSERLYS	800
2351	CATCCGAGCTTCAATATATATATATATATATATATATATATATATATAT	2400
801	LYSVALVALIASERLEUERTHREUENLYPRPRLIOUTHIPHESE	817
2401	AASGTGTGGCTTCTGTACCGCAACACTTGTAATCCGACTTGAGCAATTTCG	2450
817	TRPPTREUPROVALGLIULYSLNALALEUAGLILASUSERTILEUGL	834
4451	ACACACCTCTCCCTCCGAGAACAGCAATTCGCGGAAAAATCTATCTCTCG	2500



```

285 ArgArgValAspLeuAlaThrLeuMetAsnThrIleGluIleuGlyLeu 301
1468 AGAAAGAGUGAATCGCCACTCTCATGAACACGATCGACCAATCGGACT 1517
301 uGlyThrGluTyrThrLeuAspAsnValThrAlaValTyrArgSerGlnT 318
1518 TGGGTTTGGTACCGTTGGACACCGAACGCTGTGTATCCCTTGTGGAA 1567
318 hAlaGluGluThrGlyAspPheAlaLeuValPheValIleuValGlyThr 334
1568 GCGCTGGTCAATGAGTGGTACATGTCTGGTGGTACGTTCTGGGAGCG 1617
435 ValAlaAspAspGlyLeuLeuPheValIleuGlyIleuAsnAspThrGlnAl 351
1618 GTGGGCAACGACGAGACTCTCTTCTTCTTCTGGGAGAAATGACACCAAG 1667
351 dTyrLeuGluGluAlaIlePheAsnGlnPheAspLeuTyrGlnThrIleu 368
1668 ATATCGCGAGGAGGACATCCGAAATACGCGGACCTTACGACACTGCGC 1717
368 euGlyValTyrPheIleuGlySerProGlyTyrGlySerProGlnAspGln 384
1718 TTGGAGCATATATTAATGATGATGAGGATGAGATGAGATGAGATGAG 1767
385 hAlaAlaIleGluThrGluValArgPheGlnGysProSerAlaIleVal 401
1768 ATGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1817
401 lAlaGlnAspSerArgAsnArgGlyIleLeuProSerTyrPheTyrTyrAla 418
1818 GCGCTGAGGAGCTGCGGAGATGCGGAGATGCGGAGATGCGGAGATGCGG 1867
418 snAlaThrPheGluAsnLeuGluLeuPheProGlySerGluValTyrHis 434
1868 ATGGGAGCTTGAAGATGAGAGCTTGTCTTCTTCTTCTTCTTCTTCTTCT 1917
435 SerSerGlyValGlyGlyMetValPheGlyThrTyrPheValAlaSerAlaThr 451
1918 AGCTGTGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1967
451 rAlaLeuGluAlaGlnThrSerGlyTyrMetGlnGlyAlaThrPheAlaThr 468
1968 GCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2017
468 hAlaTyrAspProMetAsnGlyTyrGlyTyrPheGlnValProAsnVal 484
2018 TTGCTGAAGAAAGCTATGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2067
485 AlAlaLeuGlySerProGlyLysAlaIleGlnValAspValSerProAl 501
2068 GCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2117
501 aThrLeuAspGlnArgGlyAlaLeuTyrThrArgTyrTyrThrGluLeuG 518
2118 GACAAATAGAGCAACGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2167
518 lYrThrIleAlaTyrArgThrThrPheGlyTyrGlySerGlyGlySerGly 534
2168 GCGCAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2217
535 GtGlySerGlyAspAsnValAlaAspValValValGlyValGlyLeu 551
2218 GGAAGGACGCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 2267
551 uSerGlyLeuGluThrAlaIleValGlyValGlyAlaGlyLeuSerGly 568
2268 GAGGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2317
568 gValLeuGluValMetAspArgValTyrGlyTyrThrLeuSerValGln 584
2318 TCGTTCTGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2367
585 SerGlyProGlyArgThrThrIleAsnAspLeuGlyAlaAlaThrPheAs 601

```

---

```

2368 TCGGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2417
601 nAspSerAsnGlnSerGlnValSerArgLeuPheGluArgPheHisLeuG 618
2418 TCACAGCAGCAACGAGGAGATATGCAATTTCTTGTGAAGATTGATTCATG 2467
618 lGlyGluLeuGlnArgThrThrThrGlyAsnSerIleHisGlnAlaGlnAsp 634
2468 AGGCGAGAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2517
635 GtYrThrThrThrAlaProTyrGlyAspSerLeuLeuSerGluIleVal 651
2518 GGTACACACATACAGCTCTCTATGCTGAGCTCTTCTTCTTCTTCTTCTTCT 2567
651 lAlaSerAlaLeuAlaGluLeuLeuProValTyrSerGlnLeuIleGluG 668
2568 TCGCAAGCTGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2617
668 lHisSerLeuGlnAspLeuLysAlaSerProGlnAlaLysAlaGlyAsp 684
2618 AGCATAGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2667
685 SerValSerPheAlaHisTyrGlySerGluLysGluLeuAsnLeuProAlaVal 701
2668 AGCTGACCTTTCGCGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2717
701 lLeuGlyValAlaAsnGlnIleThrArgAlaLeuLeuGlyValGlnIleHis 718
2718 TCTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2767
718 lGluIleSerMetLeuPheLeuThrAspTyrIleGlySerAlaThrGly 734
2768 AGGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2817
735 LeuSerAsnIlePheSerAspGlyLysAspGlyGlyGlyTyrMetArgGly 751
2818 CTGCAATATATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2867
751 sLysThrGlyMetGlnSerIleGlyHisAlaMetSerGlySerGluLeuValP 768
2868 CAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2917
768 rGlySerValHisLeuAsnThrProValAlaGluIleGluGlnSerAla 784
2918 CAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2967
785 SerGlyCysThrValArgSerAlaSerGlyValAlaPheArgSerTyr 801
2968 TCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3017
801 sValValValSerLeuProThrThrLeuTyrProThrIleuThrPheSerP 818
3018 GCGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3067
818 rGluLeuProAlaGluIleGlnAlaLeuAlaGluAsnSerIleLeuGly 834
3068 CAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3117
835 lYrTyrSerGlyIleValPheValTyrPheLysProTyrPheArgGluG 851
3118 TACTATACAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3167
851 nGlyHisSerGlyValLeuGlnSerSerCysAspProIleSerPheAla 868
3168 AGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3217
868 rGAspThrSerIleAspValAspArgGlnTyrSerIleThrCysPheMet 884
3218 GAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3267
885 ValGlyAspProGlyArgLysTyrSerGlnIleSerTyrGlyValArgG 901

```



[illegible]









```

1240 .....GGAGAGCGCGGCTACCCAGAGCGCAATTCCTTA 1283
440 ValPheGlyThrTyr.....Pro 447
1284 CATTTCAGGAGGTTTCAGCTCGACGACGATGGATGTGTCGATTGGCGGC 1333
447 AlaLeuAlaThrPheAlaLeuGlnAlaGlnThrSerLysTyrMetGlnGly 463
1334 CGAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1383
444 AlaTTPAlaAlaPheAlaLysAsn.....ProMetAs 474
1384 GCGTCGATTCGGTTGGCGAAGATGGCCACCCGCGCGGAGCGCCCTTAC 1433
474 MetLysThrGlyTyrLysGlnValProAsnValAlaAlaLeuGlySerPro 491
1434 CTGGGCTGGCTATCTATCGGCGACGTCGACATGATTCGATTCGCGGAGG 1483
491 LysAlaAlaLeuValAspValSerProAlaThrLeuAspGlnArgCys 507
1484 GTGGCGGCG.....GGCGTGGTGTGGCGGAGCTTCGATCGCGCGCTGG 1527
508 AlaLeuTyrThrArgTyrTyrThrGlnGlyThrLeuAlaProArgTh 524
1528 GCG.....GATGGCGCGCAAGGC 1544
524 PheGlyGlyLysSerGlyLysSerGlyLysLysSerLysAsnVal 541
1545 GAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1594
541 AlaAlaAspValValValGlyValGlyLysSerGlyLysGlnThrAla 557
1595 TTGGCGATTCGGTACGTCGTCGGCGCTTCGACGCGTTTCGACGCGCA 1644
558 ArgLysValGlnAlaAlaGlyLysSerGlyLysValLeuGlnAlaMetAs 574
1645 GCGAAGATTCAGGCGCGGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGT 1694
574 PheValGlyLysThrLeuSerValGlnSerGlyProGlyArgThr 591
1695 TCGCTAGGAGGAGAAAGCTCTAAGCTTAATTCGCGCGCGCGAGCGCA 1744
591 MetLysAsnLysGlyAlaAlaThrPheAsnAspSerAsnGlnSerGlu 607
1745 GATCAACAGCCCTCGAGCTGCGTCAATCAATCAACAGCAACCAAGCGCA 1794
608 ValSerArgLeuPheGlnArgPheGlnLysGlnGlyLeuGlnArgTh 624
1795 GATTCGCAATTCGTTTAAAGATTTCATTGGAGGCGGAGCTTCGACGAGC 1844
624 ThrGlyAsnSerLeuIleGlnAlaAspGlyThrThrThrAlaP 641
1845 GACTGCAAAATTCATCATCAATCAATCAATCAATCAATCAATCAATCAAT 1894
641 ThrLysLysSerLeuLeuSerGlnValAlaSerAlaLeuValGly 657
1895 CTATATGCTACCTCTTCTGAGCGAGGCTTTCAGTGCACCTTCGCGCA 1944
658 LeuLeuProValTrpSerGlnLeuIleGlnGlnLysSerLeuGlnAspLe 674
1945 CTCTCTCTCGATGTCGTCAGCTGATCGAAGGACATACGCTTCAGAGCT 1994
674 MetLysAsnTrpGlnAlaLysArgLysAspValSerPheAlaHisTr 691
1995 GAAAGCGAAGCGCTCAAGCGAAGCGCTCAAGCGCTCAAGCGCTCAAGCT 2044
691 LysGlyGlnLysGlnLeuAsnLeuProAlaValLeuGlyValAlaAsnGln 707
2045 ACGTCAAGAGGAGCTAAAGCTTCGCTGCTGCGCTAGCAAGCGAG 2094
708 IleThrArgAlaLeuLeuGlyValGlnAlaHisGlnLysMetLeuPhe 724

```

```

2095 ATCAACGCGCTGCTGCTGAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAG 2144
724 GCGCTThrAspTyrTyrLysSerAlaThrGlyLysSerAsnLeuPheSerVa 741
2145 TGTCAAGGAGCTACATCAAGAGCTGGGACCGGCTTCAAGATATATTTCTCG 2194
741 SerLysAspGlyLysLysMetArgLysLysThrLysMetGlnSer 757
2195 ACAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2244
758 MetLysSerAlaMetSerLysGlnLeuValProGlySerValHisSerAs 774
2245 ATTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2294
774 MetProValAlaGlnLysGlnLysSerAlaSerGlyLysThrValArgS 791
2295 CAGTCCGCTGCTCAATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 2344
791 ArgAlaSerGlyValAlaPheArgSerLysLysValValValSerLeuPro 807
2345 GCGCGTGGCGCGCGCGCTTTCGAGAGCAAAAACTGCGCGCTTTCGATTCG 2394
808 ThrThrLeuTyrProThrLeuThrPheSerProProLeuThrAlaGlnLys 824
2395 ACAAGCTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 2444
824 SerAlaAlaGlnAlaAsnSerIleGlnGlyTyrTyrSerLysIleValP 841
2445 GAGAGCATTCGCGGAGAAATTCATTCGTCGCTTACATATTCATATTCATAT 2494
841 MetValTPAspLysProTTPArgGlnGlnGlyPheSerGlyValLeu 857
2495 TCGATATGGGAGCAAGCGCTGCTGGCGGAGCAAGCGCTTTCGATTCGATTC 2544
858 GlnSerSerGlyAspProLysPheAlaArgAspThrSerLeuAspVal 874
2545 CAAATGAGCTTCAGCTTCATTCATTCATTCATTCATTCATTCATTCATTC 2594
874 AspArgGlnLeuTrpSerTyrThrLysPheMetValGlyAspProGlyVal 891
2595 CGATTCGCAATTCGATTCATTCATTCATTCATTCATTCATTCATTCATTC 2644
891 LysTTPSerGlnLysSerLysGlnValArgGlnLysSerValTPAspGln 907
2645 AGTGTTCGCAAGAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2694
908 LeuArgAlaAlaTyrGlnLysAlaGlyValAlaGlyValProGlnProAlaAs 924
2695 CTGCGCGGAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2744
924 ValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 941
2745 CCGTGTCAAAATTCAGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2794
941 GlnAlaValTyrGlyLysAsnAspLeuIleThrLeuGlyLysSerAlaLeuArg 957
2795 GCGCGCTTATCGGCTCAAGCATTCATTCATTCATTCATTCATTCATTCAT 2844
958 ThrProPheLysSerValHisPheValGlyThrGlnLysLeuValTr 974
2845 ACGAGGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2894
974 LysGlyGlyTyrMetGlnGlnAlaGlnArgSerGlyGlnArgGlyAlaAlaG 991
2895 GAAAGCGATATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2944
991 LeuValValAlaSerLeuValTrpAlaAla 1000
2945 AAGTTCGCTAGCGCTGCTGCGAGAGAGCA 2973

```

seq\_name: qb\_pat:ARI45461

seq\_documentation\_block:



443 dPhetIhYrTYR.....ProVal 447  
 1901 TTTTGGAGGTTTCAGACTGAGCTGGTGTTCATTTGGCGGCC 1950  
 448 AlaSerAlaThrAlaLeuGluAlaGlnThrSerLeuTyrMetGluGlyAl 464  
 1951 GAGAGAGGACGCGCGCGCGAGCGTGGAGTGGCAATGATGATGCTGGCG 2000  
 464 dTPAlaAlaPheAlaLysAsn.....PrometAsnG 475  
 2001 GTGGAGTGGTGTGGTCAGAGATTTTACCTGGCGGAGCGCTTACCT 2050  
 475 TyrProGlyTyrPheGluValProAsnValAlaAlaLeuLysSerProGly 491  
 2051 GCGGTGGCTATTCCTACGGGCAACTGACCAATTCGTTGGCGAGCGCT 2100  
 492 LysAlaAlaLeuValAspValSerProAlaThrLeuAspGluArgCysAl 508  
 2101 GCGGAGT.....GCGGTGGTGGCGGAGCTTGGATGGGCGG 2144  
 508 dLeuTyrThrArgTyrTyrThrGluLeuGlyThrLeuAlaPheArgPhe 525  
 2145 G.....GATGGCGGTAAAGCGCG 2161  
 525 dLeuTyrTyrLysSerGlyGlyGlyLysSerGlyLysLysSerLysAspAsnVal 541  
 2162 GGGGCGAGAGTACGGGTGGAGCGTACGCGCGAGCGTACCGTAAAGCAGCT 2211  
 542 AlaAsnValValValAlaGlyAlaGlyLeuSerGlyLeuGluThrAlaArg 558  
 2212 GCGAGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2261  
 558 dLysValGluAlaAlaGlyLeuSerGlyLeuValLeuGluAlaMetAsp 575  
 2262 CAAGAGTCCAGAGCGCGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2311  
 575 TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 591  
 2312 GTGTAGGCGGAAAGACTCTGAGCGTAAATCGGCTGGCGAGGAGGAGCT 2361  
 592 dLeuAspLeuGlyValAlaAlaTyrPheAsnAspSerAsnGlnSerGlyVal 608  
 2362 ATCAATACCTTGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2411  
 608 dSerArgLeuPheGluArgPheHisLeuGluGlyGluLeuGluArgThr 625  
 2412 ATCTAGATTTGTTGAAGATTTCATTTCAGAGCGCGAGCTCCAGAGCGAG 2461  
 625 dTrpLysAsnSerLeuHisGlnAlaGlnAspGlyThrThrThrThrAlaPro 641  
 2462 GTGAAATTCATTCATCAACACAGAGCGAGAGAGCTACAGCTGGCT 2511  
 642 TyrGlyAspSerLeuLeuSerGluGluValAlaSerAlaLeuAlaLeu 658  
 2512 TATGGTACACTTCTCTGAGCTAGCTAGCTGGTGGTGGTGGTGGTGGTGG 2561  
 658 dLeuThrValTyrPheGluLeuLeuGluGluHisSerLeuGluHisLeu 675  
 2562 GTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2611  
 675 LysAlaSerProGlnAlaLysArgLeuAspSerValSerPheAlaHisTyr 691  
 2612 AGGAGAGCTTCAGAGGAGAGAGAGCTGACAGAGTGGAGCTTGGCGAGAG 2661  
 692 CysGluLysThrLeuAsnLeuProAlaValLeuGlyValAlaAsnGluLeu 708  
 2662 TGTGAG 2711  
 708 dThrAlaValAlaLeuLeuGlyValAlaHisSerThrSerMetLeuPheLeu 725  
 2712 CAG 2762  
 725 dThrAspTyrLeuLysSerAlaThrGlyLysSerAsnThrPheSerAsp 741

2762 TCACCGACTACATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2811  
 742 dLysAspGlyGlyGluTyrMetArgCysGlyThrGlyMetCysArg 758  
 2812 AAGAAAG 2861  
 758 dCysHisAlaMetSerLysGluGluValProGluValHisSerMetSer 775  
 2862 TTGGCATGCGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2911  
 775 dProValAlaAlaGluLeuGlnAsnSerAlaSerGlyTyrThrValArgSer 791  
 2912 GCGGTGGCTGAATTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2961  
 792 AlaSerGlyAlaValPheArgSerLysValValAlaSerLeuProThr 808  
 2962 GCGGTGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 811  
 808 dThrLeuTyrThrPheLeuThrPheSerProGluProAlaGluLysG 825  
 3012 AAGCTGTATCCGAGCTTGCATTTTCAGCAGCGTCTTCCGAGAGAGAGAGAG 8361  
 825 dAlaAlaGluAlaGluAsnSerLeuLeuGlyTyrTyrSerTyrLeuValPhe 841  
 3062 AAGTATTGGCGGAAATTCATTCCTGGGCTACTATAGCAGATGAGCTTTC 811  
 842 ValTyrPheLysProTyrThrPheGluGluGlyPheSerGlyValLeuGlu 858  
 3112 GTATGGCAAGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 861  
 858 dSerSerLysAspProLeuSerProAlaArgAspThrSerLeuAspValA 875  
 3162 ATCCAGCTGTACCGCCATTCATTTGCCAGATATCAGTATCCAGTACGCTGG 8211  
 875 dPheGluThrPheSerLeuThrGlySerMetValGlyAspProGlyArgLys 891  
 3212 ATCCAGCAATGCTCATTAATGTTTATGCTGGAGAGAGAGAGAGAGAGAGAG 8261  
 892 TrpSerGlnGlnSerLysGluValArgGlnLysSerValTrpAspGlnLe 908  
 3262 TGGTGGCGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8311  
 908 dArgAlaAlaTyrGluAsnAlaGlyAlaGluValProGluThrAlaAsnV 925  
 3312 GCGGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8361  
 925 dLeuGluLeuGlnThrSerLysGlnGlnThrPheGlnGluValAspProSer 941  
 3362 TGGTGGAGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8411  
 942 AlaValTyrGlyLeuAsnAspLeuLeuThrLeuGlySerAlaLeuAlaThr 958  
 3412 GCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8461  
 958 dProPheLysSerValHisPheValGlyThrGluThrSerLysValTyrP 975  
 3462 GCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8511  
 975 LysGlyTyrMetGluValAlaGluLeuArgGlySerLysArgGlyValAlaGlu 991  
 3512 AAGGTATATGAG 8561  
 992 ValValAlaSerLeuValProAlaAla 1000  
 3562 GTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 8588  
 seq\_name: gb\_pat:AK145488  
 seq\_documentation\_block: 3591 bp CNA linear Feb 08 AUG 2001  
 Locus: AK145488  
 DEFINITION Sequence 30 from patient US 6211445.  
 ACCESSION AK145488







ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 1947)  
AUTHORS Puntieri, J., Maddox, J. E., Ford, T. A., and Wang, X.  
TITLE Fumonisin detoxification: compositions and methods  
JOURNAL Patent: US 6,229,921-A 15-09-MAY 2001  
FEATURES  
source 1. 1937  
protein 448 a 525 c 487 g 477 e

alignment\_scores:  
Quality: 3440.00 Length: 512  
Ratio: 5.068 Gaps: 3  
Percent Similarity: 96.863 Percent Identity: 94.096

alignment\_block:  
US-09-658-835C-25 X AK150992 ..

Align seg 1/1 to: AK150992 from: 1 to: 1947

```

1  MetAlaAsnLysHisLeuSerLeuSerLeuPheLeuValLeuLeuIle 17
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
153 ATGGCGTGAAGCTACATCTCTCTTGGCTGTACCTGCTTTGGCGCAT 202
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
17  usor  AAsortLeuAlaSerGlyAlaPheThrValIleGlyAlaSPA 32
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
203 TGGCTTGGTGGCAATAGCGGGTGGCTGGCTGCTACCTGCACACATGATC 252
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
32  IacIyMeValValIleThrThrThrValAlaProGlyThrThrAlaThr 48
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
253 CTGGATAGCGTGGTGGACAGACTACTGCTGCGCCGCGACGCGCGCCG 302
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
49  ValSerGlyPheLeuGlyValProPheAlaAlaSerProThrArgPheAl 65
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
303 GTCCAGCAAGTCTTGCGGTCTTGGCGCGCTCCGACACGATTTCC 352
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
65  AProThrArgPheValProIlePheSerThrProLeuGlnAlaThrAlaT 82
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
353 GCGCTGACTGCTGGCTGGCTTGGTCAAGCGCTTGGCAACGCTGCAT 402
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
82  YrGlyProAlaCysProGlnGlnPheAsnIleProGlnGlnLeuArgLys 98
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
403 ATGGTCCAGCAAGTCTGCAACATTCATTAACCCGCAACGACGCGGAC 452
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
99  IleThrMetAlaIlePheAsnIleThrProProSerAlaGlyIleuSerG 115
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
453 ATTAGGATGCTGCTGATTAATACACGCGCGCGCACTGCTGAAGTGA 502
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
115  UAspGlySerArgTyrGlnGlyTyrValProGlyThrGlnAspThrArg 122
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
503 GGACTGGCTGAAGCTCAACATCTAGCTGCGGAGAACGACACACGAAACA 552
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
137  YrAlaValIleValIlePheIleThrIleGlyAlaGlyIleThrIleAsn 148
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
553 AAGCGCTATGCTGCTGATATACGCTGCAACGCTGCAATACGCTGCAAT 602
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
149  SerPheIleLeuIleThrGlyAlaSerPheAlaAlaLeuGlnAspValIle 165
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
603 TCATTCACCTTTTCACGCGCGCTAGCTTTGGACGCGCATCGGATGTCAT 652
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
165  AlaValIleIleIleAsnIleThrAsnIleLeuGlyIlePheProAlaIle 182
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
653 CGTGTGACGATTCACCTACAAACGACATCTGGGGTCTCCCTGCTGGCG 702
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
182  rGlnIleuPheIleThrIleArgAsnLeuGlyIleLeuAspGlnArgPhe 198
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
703 CTGAGCTTCACAAACACGAGAAATCTGGGGTCTCTAGACCAAGGCTTT 752
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
199  AlaLeuAspIleValIleArgAsnIleAlaIlePheIleGlyIleAspPro 215
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

753  GCTTGGATTTGGTATACAGCGCAATCGCAAGGCTTGGGGGCTGATCTGG 802
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
215  GlyValIleThrIlePheGlyIleHisSerAlaGlyIleArgSerValAspVal 232
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
803  AAGGCTACAAATATATATATATATATATATATATATATATATATATAT 852
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
232  LeuLeuThrSerMetProHisAsnProProPheArgAlaAlaIleMetGly 248
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
853  TCTTATGCTATATATATATATATATATATATATATATATATATATAT 902
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
249  SerGlyValAlaAsnIleAsnPheProGlyGlyAspLeuSerIleProTr 265
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
903  TGGTGTGATGATACATACAAATTCGCGCAAGGATTTGCTGCAACCTGG 952
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
265  PAsnIleThrValGlnAlaLeuAsnGlyThrThrSerLeuAspIleLeuS 282
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
953  GAAATACATGCTTAAAGCTGTCAATCTGATACACATATGATGATGGA 1002
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
282  GCGYSMeLAlaGlyValAlaIleValIleThrIleuMeLAsnThrIleGly 298
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1003  GTTGTATATATATATATATATATATATATATATATATATATATATAT 1052
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
299  LeuGlyLeuGlyIlePheGlyIleThrIleuAspAspValThrAlaValTyr 315
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1053  CTGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1102
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
315  GSerGlyIleThrAlaArgThrThrGlyAspIleAlaArgValProValLeu 332
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1103  TTTTCAAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAA 1152
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
332  AIGlyThrValAlaAsnAspGlyLeuLeuPheValLeuGlyIleuAsnAsp 348
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1153  TCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1202
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
349  ThrGlnAlaTyrIleProGlyAlaIlePheAsnGlyIleProAspIleu 365
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1203  ACCCAATCATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
365  rThrLeuLeuGlyValaTyrProIleGlySerProGlyIleGlyIleSerPro 382
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1253  GATCTCTCTTGGTAAATATCTTATGCTTATGCTTATGCTTATGCTTAT 1302
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
382  LAspGlnIleAlaAlaIleGlyIleuGlnValArgPheGlnCysProSer 398
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1303  AAGATCAGATTCGCGCATTCAGACCGAGGTAAAGATTCAGATTCAGCT 1352
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
399  AlaIleValAlaIleAspSerArgAsnArgGlyIleProSerTyrAlaGly 415
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1353  GCTTATCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 1402
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
415  TTYTyrAsnAlaIleThrPheGlyAsnIleuGlnLeuPheProGlySerGly 432
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1403  GCTGCTTAAATGCTGCTTGGGAAATCTGAAAGCTTCTTCTTCTTCTT 1452
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
432  aTyrHisSerSerGlyValaGlyMetValPheGlyThrTyrProValAla 448
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1453  TGTACATACATCTGTAAGTGGGATGCTTGGTGGCTGATCTGCTGCTG 1502
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
449  SerAlaIleThrAlaLeuGlnAlaGlnThrSerLysTyrMetClnGlyAla 465
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1503  ACTTATATATATATATATATATATATATATATATATATATATATAT 1552
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
465  PAlaAlaPheAlaIleLysAsnProMetAsnIleProGlyTyrGlyGlnVal 482
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1553  GCGGCTCTTCTGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTT 1602
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
482  rGAsnValAlaAlaIleGlySerProGlyLysAlaIleGlnValaAspVal 498
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1603  TAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1652
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
499  SerProAlaIleThrIleAspGlnArgCysAlaIleuTyrIleArgTyrThr 515
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
1653  TCTCCAGCGACATATACACGATATGCTTGTATACACCATATATATATAC 1702
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

```







```

568 LeuValIleuGluAlaMetAspArgValIleGlyIleThrLeuSerValGlu 584
592 CUGGTCCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
584 AsnGlyIleGlyArgIleThrIleAspIleuIleValIleIleThrLeu 601
592 ATGGATTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
601 snAspSerAsnIleSerValIleSerAlaIleuPheGluIleArgPheIleu 617
602 AGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 651
618 GluIleGlyIleuIleuIleArgThrIleGlyAsnSerIleHisGluAlaGlu 634
652 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701
634 polyThrIleThrIleAlaProGlyIleIleSerIleuLeuSerGluIleu 651
702 GGTATCAAGGACTTAAGTCTGATGATGATGATGATGATGATGATGATGATGAT 751
651 dAlaSerAlaIleuAlaIleuIleuProValIleIleSerIleuIleuIle 667
752 TTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
668 GluHisSerIleuIleAspIleuGlyAlaSerIleuIleAlaIleuIleu 684
802 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 851
684 psrValIleuPheAlaHisIleIleGlyIleuIleuIleuIleuIleuIle 701
952 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901
701 dIleuIleuValAlaAsnIleuIleuArgAlaIleuIleuIleuIleuIle 717
902 TTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 951
718 HisGluIleuSerIleuPheIleuIleuIleuIleuIleuIleuIleu 734
952 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1001
734 yIleuSerAsnIlePheSerAspIleuIleuIleuIleuIleuIleuIle 751
1002 TTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1051
751 yIleuSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 767
1052 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
768 ProGlySerValIleIleuAsnIleuIleuIleuIleuIleuIleuIle 784
1102 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151
784 aserIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 801
1152 ATGGATTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1201
801 yValValValSerIleuProIleuIleuIleuIleuIleuIleuIleu 817
1202 ATGGATTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1251
818 ProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 831
1252 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1301
834 yIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 851
1302 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1351
851 IleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 867
1352 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1401

```

```

868 ArgAspIleuSerIleAspValAspArgGluIleIleThrIleuPheIleu 884
1402 AGAGATACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCA 1451
884 ValGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 901
1452 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1501
901 IleuSerValIleuPheIleuIleuIleuIleuIleuIleuIleuIleu 917
1502 AAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551
918 GluValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 934
1552 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1601
934 nIleuPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 951
1602 GATTTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1651
951 IleuGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 967
1652 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1701
968 ThrGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 984
1702 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1751
984 rGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1000
1752 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800

seq_name: qb-pat:AR145462
seq_documentation_block:
LOCUS AR145462 2490 bp DNA linear 161 08 Apr 2001
DEFINITION Sequence 32 from patent US 6211434.
ACCESSION AR145462
VERSION AR145462.1 GI:15107329
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2490)
AUTHORS Duivick,J.P., Gilliam,J.T. and Maddox,J.R.
TITLE Anticancer agent and methods of use
FEATURES
location/Qualifiers
1..2490
BASE COUNT 625 a 623 c 653 g 589 t
ORIGIN
alignment_scores:
Quality: 2490.00 Length: 651
Ratio: 4.369 Gaps: 8
Percent Similarity: 84.025 Percent Identity: 75.883

alignment_block:
US 09 658 835c 25 A AR145462 ..
At:qin seq 1/1 to AR145462 from 1 to 2490

355 GluAlaIleuProAsnGluIleuAspIleuIleuIleuIleuIleuIleu 371
638 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 707
731 ProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 746
754 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 764

```

[illegible]







```

767 ValProGlySerValHisLeuAsnThrProValAlaGlnIleGluGluGluSer 783
|||||
1375 GTTCAGGCTCAAGACCTCAAGAGCGCGGCGCTGAAATTCAGACAGCTC 1424
783 LAlaSerGlyCysThrValArgSerAlaSerGlyValAlaValPheArgSerL 800
|||||
1425 GGCATCCGCGCGCTCAAGTACGATACGCGCTCGACGCGCGCTGCGACAGTA 1474
800 yslGlyValValSerLeuProThrIleLeuTyrProThrLeuThrPhe 816
|||||
1475 AAAAGTGGTGGTTGGTACACGAAACGCTTCAACCTGACCTTGACATTT 1524
817 SerProGluLeuProAlaGlnIleGlySerIleAlaGluAlaGluHisSerIleG 833
|||||
1525 TCACTACCTCTTCCGCGCAAGAAATCAAGATTCGCGCAAAATTTATGCT 1574
833 GCLYGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 950
|||||
1575 GAGCTACATACAAATAAGCTTGGTATACGTAACGCGCGCTGGCGCGCG 1624
850 IuclucIyPheSerThrValLeuGlnGlySerCysAspProIleSerPhe 866
|||||
1625 AACAAAGCTTCTCGACGCGCTCTCAATGACGCTGACGCGCAATCTCATTT 1674
867 AlaAlaGlySerThrSerIleAspValAspArgIleThrPheSerIleThrCysPh 883
|||||
1675 GCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1724
883 GMeValIleIAspProGlyArgLysTrpSerGlnIleSerIleValAla 900
|||||
1725 CATGGCGGAGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1774
900 PGLIYSSSPVALLRPASPGIAGLQATGALALGTYRGLUASGLAGLY 916
|||||
1775 GACAAAGATTCGTCTGGGAGCAACTCCGCGGAGCTACGACAAATCCCGCG 1824
917 AlacIlnValProGlnProAlaAsnValLeuGlnIleGlnIleTrpSerLysG 933
|||||
1825 GCCCAAGTCGTCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1874
933 nclIutyPheGlnIleValProSerAlaValItyrGlyLeuAsnAspLeu 950
|||||
1875 GCAAGATTTCCAGAGAGGCTCCGACGCGCGCTATCGGCTCAACAGATCTCA 1924
950 lctThrLeuItySerAlaLeuArgGlnProPheItyLysSerValHisPheVal 966
|||||
1925 TCACACTGGGCTCGAGGCTCAGAGAGCGCGTTCACAGATGTTTATTTGCTT 1974
967 GlyThrGlnThrSerLeuValTrpLysGlyItyMetGlnItyAlaIleArg 983
|||||
1975 GAAATCGAGAGGCTCTTAACTTCGAAAGAGGATATCGAAGCGGTATAGG 2024
983 qSerGlyItyIleArgGlyAlaIleIleValValAlaSerIleArgValIleAla 1000
|||||
2025 ATCCAGATCAAGGAGGCTCCGAGCAAAATTCGCGCAAGCTCCGCGCGAG 2074
1000 Ld 1000
||
2075 CA 2076

```







1601 GTCAGAGGACCAAGATAACCTTGGACAGTGGTACGCTGGC 1650  
 551 LPSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTGY 577  
 1651 TTAGACGAGTTTGGACAGGCAACAAATTCACAGGCGCGGTGGTGG 1700  
 567 SLQVVALLEUQUALMCAASFAVAIGLYGILYLSLTHLPSSTVALG 584  
 1701 GCTGCTTCTTAAAGGATGATGATGATGATGATGATGATGATGATG 1750  
 584 LPSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTGY 606  
 1751 AATCGAGTTCGAGGACGACGACGACGACGACGACGACGACGACGAC 1800  
 601 ASASPSSTASSTGSLPSSTGSLPSSTGSLPSSTGSLPSSTGSLPSST 617  
 1801 AATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1850  
 617 GTHGVLPLGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 634  
 1851 GAGAGGACGACGACGACGACGACGACGACGACGACGACGACGACG 1900  
 634 SPGLYTHRTHTHTALAPROTYSIYASPSSTGSLPSSTGSLPSST 650  
 1901 AAGATACGACGACGACGACGACGACGACGACGACGACGACGACG 1950  
 651 VALIASSRTALALALALALALALALALALALALALALALALALAL 667  
 1951 GTTGACATGACGACGACGACGACGACGACGACGACGACGACGACG 2000  
 667 GCTHSSSTGSLPSSTGSLPSSTGSLPSSTGSLPSSTGSLPSSTGSL 684  
 2001 AAGAGATACGACGACGACGACGACGACGACGACGACGACGACGAC 2050  
 684 SPSTVALPSSTGSLPSSTGSLPSSTGSLPSSTGSLPSSTGSLPSST 700  
 2051 ACAGTGTGACGACGACGACGACGACGACGACGACGACGACGACG 2100  
 701 VALLEUUGLYVALALASGTHRTHTHTALAPROTYSIYASPSSTG 717  
 2101 GTTGGACGACGACGACGACGACGACGACGACGACGACGACGAC 2150  
 717 GCTHSSSTGSLPSSTGSLPSSTGSLPSSTGSLPSSTGSLPSSTGSL 734  
 2151 GAGAGATACGACGACGACGACGACGACGACGACGACGACGACGAC 2200  
 734 LYSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSST 750  
 2201 GCTGCTTCTTAAAGGATGATGATGATGATGATGATGATGATGATG 2250  
 751 GYSLGSLTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 767  
 2251 TGCAGAAAGAGATGAGATGATGATGATGATGATGATGATGATGATG 2300  
 767 LPSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTG 784  
 2301 TGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2350  
 784 LPSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTG 800  
 2351 GATTCGAGTTCGAGGACGACGACGACGACGACGACGACGACGACG 2400  
 801 LYSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTG 817  
 2401 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2450  
 817 LPSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTG 834  
 2451 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2500  
 834 LYSSTGGYLAUUGUUTUAATAPLYSAIIGHUAALAGLYLPSSTG 850  
 2501 GCTGCTTCTTAAAGGATGATGATGATGATGATGATGATGATGATG 2550

851 GlnGlyPheSerGlyValLeuGlnSerCysCysAspProIleSerPheAl 867  
 2551 CAAGGTTCTGGGAGCTGCTGCATGAGCTGTGAGCGCATCTTCATTTGG 2600  
 867 ArgAspThrSerIleAspValAspArgGlnTrpSerIleThrCysPheM 884  
 2601 CACGATATCCACCATCCACGATCCGATCCGATCCGATCCGATCCGAT 2650  
 884 eValGlyAspPheGlyValGlyTrpSerGlnGlnSerGlnValAla 900  
 2651 TGGTGGACGACGACGACGACGACGACGACGACGACGACGACGACG 2700  
 901 GlnYSerValTrpAspGlnLeuArgAlaAlaTyroValAsnAlaCyl 917  
 2701 CAAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2750  
 917 GlnValProGlnProAlaAsnValLeuGlnIleGlnTrpSerGlyGln 934  
 2751 CCAAGTGGTGAAGGCGAGGCAAGCTGCTGCAAAATGCAATGCTGCA 2800  
 934 IntYrPheGlnGlyAlaProSerAlaValTyroGlyLeuAsnAspLeu 950  
 2801 AGTATTTCGAGGAGGCTGCGACGCGCGCTGATGGCTGAGACGATCT 2850  
 951 ThrLeuGlySerAlaLeuArgThrProPheLysSerValHisPheVal 967  
 2851 AACTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2900  
 967 YrPheGlnThrSerLeuValTrpYrGlyTrpMetCylGlyAlaValArg 984  
 2901 AACGATACGCTTTTACTTTTGCAGAGTATATATATATATATATATAT 2950  
 984 eArgGlyAlaGlyValAlaGlnValValAlaSerLeuValProAlaVal 1000  
 2951 GCGGTGACGACGACGACGACGACGACGACGACGACGACGACGACG 3000

seq\_documentation\_block:  
 ID AA260642 standard; DNA; 3003 BP.  
 XX  
 AC AA260642;  
 XX  
 DT 16-MAY-2000 (first entry)  
 XX  
 DE rRNA encoding a fusion of aminoglycosyl aminocyclitolymonisin esterase.  
 XX  
 KW Aminoglycosyl aminocyclitolymonisin dephosphorylase; aminoglycosyl  
 KW plant fungal invasion; hydrogen peroxide; fusarium; fungus;  
 KW fumonisin esterase; ss.  
 XX  
 OS Synthesis.  
 OS Exophiala spinifera.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 1..3003  
 FT /tag= a  
 FT /feature= "aminoglycosyl aminocyclitolymonisin  
 FT /feature= "esterase fusion"  
 FT sig\_peptide 1..72  
 FT /tag= b  
 FT /note= "partial alpha amylase signal sequence"  
 FT mat\_peptide 73..1575  
 FT /tag= c  
 FT /product= "fumonisin esterase"  
 FT misc\_feature 1576..1611  
 FT /tag= d  
 FT /note= "spacer sequence"  
 FT mal\_peptide 1612..3000  
 FT /tag= e  
 FT /product= "aminoglycosyl amine oxidase"  
 FT /note= "the protein has an amino terminal lys



1151 AGAATGGCCGATTAGACGAGGAAAGATTCAGCTGCTCTCCGATC 1200  
 401 ValAlaGlnAspSerArgAsnArgGlyIleProSerTrpArgTyrTyr 417  
 1201 GAGGCTGAGACTGCGGAATCGAGATCGCTCTCTGACGCTACTACTA 1250  
 417 PAsnAlaIlePheGluAsnGlnGlnLeuPheProGlySerGlnValTyr 434  
 1251 CAATGGACCTTCAATCTGAGGCTTTCCCTGGGTCGCAAGTGAAC 1300  
 434 IsserSerGluValIleGlyMetValPheGlyThrThrProValAlaSerAla 450  
 1301 AACGCTCTGAGTCGCGATGCTCTTGGCAGCATTCCTTGGCAAGTGGC 1350  
 451 ThrAlaLeuGluAlaGlnThrSerLysIleGlyMetGlnGlyAlaTrpAla 467  
 1351 ACCGCTTCAGAGTCAGACGACAAATACATGTAAGCTGCGCGCGG 1400  
 467 aPheAlaIleAsnProMetAspGlyProGlyTrpLysGlnValProAsnY 484  
 1401 CTTCGCAAAACCCCATGATGGGCTGGGAGAACAGATGCCAATG 1450  
 484 AlaAlaAlaLeuGlySerProGlyLysAlaIleGlnValAlaSerPro 500  
 1451 TCGAGGCTTGGCTTCACAGGCAAAAGCCATCCAGGTTGAGGCTGCGCA 1500  
 501 AlaThrIleAspGlnArgGlyAlaLeuTyrThrArgTyrThrGlnG 517  
 1501 GCGACATACGCAACGATGCTGCTGATACGAGGCTATATTAACAGGT 1550  
 517 uGlyThrIleAlaProArgThrPheThrLysGlyLysSerGlyGlyPro 534  
 1551 GAGCAATATGCTGACAGCATTTCTGGAGAGCAAGAGGAGAAAGGAGG 1600  
 534 LysGlyLysSerLysAspAspValAlaAspValValValAlaGly 550  
 1601 GCGAAATATCAAAATCAAGCTGGGAGCTGCTACGCTGGGCTGGGCT 1650  
 551 LeuSerGlyProGlnThrAlaArgLysValGluAlaGlnGlyLeuSer 567  
 1651 TTAGGCGGCTTGGAGAGGCAAGCAAGATCGAGGCGCGGCTGTGCTG 1700  
 567 sLeuValLeuGluAlaMetAspArgValGlyLysIleLeuSerValG 584  
 1701 CCGCTTTTGGATGCAATGATGCTGATAGGAGCAAGATCTTGAAGGCTAC 1750  
 584 InsSerGlyProGlyArgThrThrIleAsnAspLeuGlyAlaAlaTrp 600  
 1751 AATCGGCTCCGAGAGAGACATCAACGATGCTGAGCTGGCGGATC 1800  
 601 AsnAspSerAsnGlnSerGluValSerArgPhePheGluArgPheHis 617  
 1801 AATGACAGCAACCAAAAGCAAGATGATGATGTTGTTGAAGATTCATT 1850  
 617 uCUGlyGlnLeuGlnArgThrThrGlnLysSerIleHisGlnAlaGln 634  
 1851 GGAAGGAGAGCTCCAGAGAGAGCTGGAATTCATCCATCAAGACCAAG 1900  
 634 sPheIleThrThrThrAlaProTyrGlyAspSerLeuLeuSerGluGln 650  
 1901 ATGGTAAATATATATATATATATATATATATATATATATATATAT 1950  
 651 ValAlaSerAlaLeuAlaGlnLeuGlnProValIlePheSerGlnLeuIle 667  
 1951 GTTGCAGATGACCTTGGCAAGTCTCGGATATATATATATATATATAT 2000  
 667 uGlnHisSerLeuGlnAspLeuLysAlaSerProGlnAlaLysArgLeu 684  
 2001 AGAGCATACCTTCAAAATATCAAGAGGAGAGGCTTCAAGGCAAAAGG 2050  
 684 sPheValSerPheAlaHisIleTyrGlyLysGlnLeuAsnLeuProAla 700  
 2051 ACAGTGTACCTTGGCGACATAGCTGAGAGAGGACCAAACTTGGCTGCT 2100

701 ValLeuGlyValAlaAsnGlnIleThrArgAlaLeuLeuGlyValGluAl 717  
 2101 GTTTCAGAGATATCAAAATATATATATATATATATATATATATATAT 2150  
 717 aHisGlnIleSerMetLeuPheLeuThrAspTyrIleLysSerAlaTrpG 734  
 2151 CCACGAGATACATAGCTTTTCTTCTACCGATATATATCAAGATGCTAC 2200  
 734 LLeuSerAsnIlePheSerAspLysLysAspGlyGlnIleTyrMetArg 750  
 2201 CUCCTGACATATATATCTGCGACAAAGAAACGCGCGGCTACATATAC 2250  
 751 CysLysThrGlyMetGlnSerTLeuGlySerAlaAsnSerTyrGlnLeu 767  
 2251 TCAAAATCAATATATATATATATATATATATATATATATATATATAT 2300  
 767 lProGlySerValHisLeuAsnThrProValAlaGlnIleGlnGlnSer 784  
 2301 TCCASCTTCATGATATATATATATATATATATATATATATATATATAT 2350  
 784 lAsSerGlyCysThrValArgSerAlaSerGlyAlaValIlePheArgSer 800  
 2351 CACCGAGCTGATATATATATATATATATATATATATATATATATATAT 2400  
 801 LysValValValSerLeuProThrThrLeuTyrProThrLeuThrPhe 817  
 2401 AATCTGAGCTTCCTGATATATATATATATATATATATATATATATATAT 2450  
 817 rProProLeuProAlaGlnLysGlnAlaLeuAlaGlnAsnSerIleLeuG 834  
 2451 ATCTATCTCTGCTGCTGAGAAAGCAAGCATATGAGGAGAAATATCAT 2500  
 834 LTTTyrSerLysIleValPheValTrpAspLysProTrpTrpArgGln 850  
 2501 GCTACTATATAGCAAGATATCTGCTATATGAGCAAAAGCTGAGGAGCA 2550  
 851 GlnGlyPheSerGlyValIleGlnIleSerSerGlyAspProIleSerPhe 867  
 2551 TAAAGCTTCTGAGAGCTGATATATATATATATATATATATATATATAT 2600  
 867 aArgAspThrSerIleAspValAspArgGlnTrpSerIleThrCysPhe 884  
 2601 CAGAT 2650  
 884 eValGlyAspProGlyArgLysTrpSerGlnGlnSerTyrGlnValArg 900  
 2651 TATCTGAGAT 2700  
 901 GlnLysSerValTrpAspGlnLeuArgAlaIleArgLysAsnAlaGlyAl 917  
 2701 TAAAT 2750  
 917 aGlnValProGluProAlaAsnValLeuGlnIleGlnTrpSerTyrGlnG 934  
 2751 CCAAGTCCCAATAGCTGAGCAAGGATATGAGAAATATGAGTGGATATAG 2800  
 934 lIleTyrPheGlnGlyAlaProSerAlaValTyrGlyLeuAsnAspLeuIle 950  
 2801 AGATATTCGCAATATATATATATATATATATATATATATATATATATAT 2850  
 951 ThrLeuGlySerAlaLeuArgThrProPheLysSerValHisPheValG 967  
 2851 ACATCTGCTTTCGCTTCAAGACCGGCTTCAAGATATGCTTATTTCTGCT 2900  
 967 yThrLThrSerLeuValTrpLysGlyTyrMetGlnGlyAlaIleArgS 984  
 2901 AACGCAATATCTTTAGTTTATGAAATATATATATATATATATATATATAT 2950  
 984 eGlyGlnArgGlyAlaAlaGlnValValAlaSerLeuValProAlaAla 1000  
 2951 TGGGTCAT 3000

```

sequence: 5' ATG GGT GAT GGG TTT GAG CCG CCG GGC CCG CCG AAT TGA AGT
seq_documentation_block:
ID      AA058916 standard; cDNA, 3618 nt.
XX
XX      AA058916:
XX
XX      27 MAY 2000 (first entry)
XX
XX      DNA encoding fumonisin esterase-aminol polyol amine oxidase fusion.
XX
XX      Bacillus pumilus strain AYO+ expressing esterase, mycotoxin,
XX      transglutaminase, plant detoxification, animal feed silage,
XX      glutathione S-transferase alpha amyrase; signal peptide; maize;
XX      ss.
XX
XX      Chimera - Schistosoma japonicum.
XX      Chimeric - Eucophila spalliera.
XX
XX      Key Location/Qualifiers
XX      sta_peptide 1..687
XX      /*tag= a
XX      /product "GST _ polylinker"
XX      mat_peptide 677..3615
XX      /*tag= b
XX      /product "Fumonisin esterase-K trypsin fusion"
XX      misc_feature 2227..2229
XX      /*tag= c
XX      /note="extra lysine codon"
XX
XX      W:200004159 At.
XX
XX      27 JAN 2000.
XX
XX      08-JUN-1999: 99WD-US15454.
XX
XX      15 JUN-1999: 99US-0097049.
XX      21-MAY-1999: 99US-013591.
XX
XX      (PION-) PIONEER HI-BRED INT INC.
XX      (CURA-) CURACON CORP.
XX
XX      Duvick JP, Gilliam JT, Maddox JR;
XX
XX      WP1: 2000-102425/15.
XX      OR
XX      P-PDBR; AA058916.
XX
XX      New isolated polynucleotides, polypeptides useful for detecting and
XX      defining fumonisin or structurally related mycotoxin in processed
XX      grain or in silage -
XX
XX      Example 13; Page 115-120, 154pp; English.
XX
XX      The present sequence is that of a polynucleotide encoding a
XX      fusion protein (see AA058916) composed of glutathione S transferase
XX      and Eucophila spalliera fumonisin esterase (ESPL) joined via a
XX      peptide linker to K-trypsin (see AY5890), a truncated, put still
XX      functional, amino polyol amine oxidase of B. spalliera. The
XX      construct was designed for expression in Escherichia coli as a hybrid
XX      protein having ESPL and APAO activity. The invention provides APAO
XX      polynucleotides (see AA05893-8/7) and polypeptides (see AA05899-95) of
XX      B. spalliera and Fumonidella atroviridis. The polynucleotides
XX      are used to transform plant cells normally susceptible to Fusarium
XX      or other toxin-producing fungus infection. Also provided are
XX      methods for expressing APAO, optionally as a fusion protein with
XX      immunohistochemical, in transgenic plants, prokaryotic and non-plant
XX      eukaryotic systems. Methods for detection of grain, grain
XX      processing, silage, food crops and in animal feed and rumen
XX      microorganisms are also disclosed.
XX
XX      Sequence 3618 nt, 884 A; 307 G; 873 C; 873 T; 0 others;

```

[illegible]



251 IAlAsnIYrAsnIheProlYsAlYAspIeSerGlnProIrfAsnIhT 268  
 1368 GGGTAATTAACATTCGGGCAAGGAGCAATTCGCAAACTTCGCAAGCA 1417  
 478 hValGlnAlaLeuAspYsThhSerIleAspIleLeuSerGlyMet 284  
 1418 CIGTTCAGAGTTCACCTGATACCAATATCCAAATGCTGAGTGTATG 1467  
 285 ArrAaValAspIeAlaIhLeuMetAsnIhTLeuGlnLeuGlyLe 401  
 1478 AGAAATATGAAATTCGGCAATTCATTAACAGAAATTCAGCAATCGAA 1517  
 401 uGlyPhehIhTyrIhLeuAspAsnValIhAlaValIhYrAsnIe 418  
 1518 TGGGATTCAGTAACGTTGTAACAGTAACGCTGCTGACCTTCGAAA 1567  
 418 hValAaGThhIhGlyAspIleAlaIhValProValLeuValGlyTh 444  
 1568 GGGATTCAGCAATTCGGTAATTCCTGCTGACTTCTCTGTCGGAGC 1617  
 445 ValAlaAsnAspIyLeuLeuPheValLeuGlyLeuAsnAspIhIhAl 451  
 1618 GTGGCAAGAGAGACTGCTGCTGCTGCGGCAATTCAGAACCAAGC 1667  
 451 aTyhLeuGlnAlaIhIhPheAsnIhProAspIeYrIhIhLeuL 468  
 1668 ATATTCAGAGAAATTCAGTAACAGCTGCTTTCAGCAATCTCTG 1717  
 468 euChIAlaIhTyrIhLeuYsEProGlyIhLeuYsEProGlyAspIh 484  
 1718 TTGAGCAATACGATTCGATTCAGCAATTCAGATTCGCTGCAATCG 1767  
 485 ILeuAlaIhIhLeuIhTyrIhValIhIhPheIhYsEProSerAlaIhLeu 401  
 1768 ATTCAGCAATTCAGTAACAGTAATTCAGATTCGCTGCAATCG 1817  
 401 IAlaIhAspSerAlaAsnAspIyIhIhPheSerIhPArATyTyrIh 418  
 1818 GGTTCAGCAATTCAGTAACAGTAATTCGCTGCTGCAATTCGCTG 1867  
 418 sValAlaIhPheAlaAsnLeuIhLeuPheProGlySerGlnValIhThS 444  
 1868 ATTCAGCAATTCAGTAATTCGCTGCTGCAATTCGCTGCAATTC 1917  
 445 SerSerGlnValIhMetValPheGlyIhTyrIhProValAlaSerAlaTh 451  
 1918 AGCTTCAGTAATTCAGTAATTCGCTGCTGCAATTCGCTGCAAT 1967  
 451 FAlaLeuAlaIhIhIhSerIhYsEProGlnGlyAlaIhFAlaIhAP 468  
 1968 GGGCTTCAGCAATTCAGTAATTCGCTGCTGCAATTCGCTGCAAT 2017  
 468 hValIhIhAspIhProMetAsnIhIhProGlyIhIhValIhProAsnVal 484  
 2018 TTGCAAAATTCAGTAATTCGCTGCTGCAATTCGCTGCAATTC 2067  
 485 AlAlaLeuGlySerProGlyIhValIhIhLeuValAspValSerProAl 501  
 2068 GGGGATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2117  
 501 aThIhIhAspIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 518  
 2118 GATTAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2167  
 518 IYrIhIhLeuIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 544  
 2168 GCAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2217  
 545 GYhIh 551  
 2218 GATTAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2267

551 uSerGlyLeuGlnIhThAlaIhGlyValIhIhIhIhIhIhIhIhIh 568  
 2268 GAGCGCTTCGAGACGCGCAAGCAATTCAGCGCGCGCGCTGCTGCTG 2317  
 568 euValLeuGlnAlaMetAspArgValIhIhIhIhIhIhIhIhIhIh 584  
 2418 TCGTTCATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2467  
 585 SerGlyProGlyIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 601  
 2468 AGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2517  
 601 hAspSerAsnIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 618  
 2418 TGAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2467  
 618 IuGlyLeuLeuGlnIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 634  
 2468 AGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2517  
 635 GlyIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 651  
 2518 GATTAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2567  
 651 IAlaSerAlaIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 668  
 2568 TGGAGTTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2617  
 668 IuIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 684  
 2618 AGATTAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2667  
 685 SerValSerPheAlaIhIhIhIhIhIhIhIhIhIhIhIhIhIh 701  
 2668 AGTTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2717  
 701 ILeuGlyValAlaIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 718  
 2718 TCG 2767  
 718 IAspIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 734  
 2768 AGGATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2817  
 735 LeuSerAspIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 751  
 2818 GTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2867  
 751 SGTTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2917  
 2868 GAAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2967  
 768 ProGlySerValIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 784  
 2918 GAGGTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 2967  
 785 SerGlySerValIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 801  
 2968 TCGGTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 3017  
 801 sValIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 818  
 3018 GGGGTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 3067  
 818 ProProLeuProAlaIhIhIhIhIhIhIhIhIhIhIhIhIhIh 834  
 3068 CACTTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 3117  
 835 TyrTyrSerIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 851  
 3118 TACTTAATTCAGTAATTCGCTGCTGCAATTCGCTGCTGCAATTC 3167  
 851 nGlyIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIhIh 868





[illegible]

	seq_name: "Fumonisin-esterase-amylase fusion - 13"	Accession: AA758404
	seq_documentation_block:	
XX	ID	AA758404 standard; CDNA; 2976 bp.
XX	AA758404:	
XX	27 MAY 2000 (first entry)	
XX	DNA encoding Fumonisin esterase amylase polypeptide oxidase fusion.	
XX	Fusion Polypeptide sequence: AAY58915 + K11APV0 (see AAY58915) joined via a peptide linker to K11APV0 (see AAY58915), a truncated, but functional, amino polyol amine oxidase of Exophiala spinifera. The construct was designed for expression and secretion in plants of a hybrid Petunia having BSV1 and APDO activity. The invention provides APDO polynucleotides (see AA758383-87) and polypeptides (see AA758900-05) of P. spinifera and uninfectious derivatives. The polynucleotides are used to transform plant cells normally susceptible to Fusarium or other toxin-producing fungus infection. Also provided are methods for expressing APDO, optionally as a fusion protein with fumonisin esterase, in transgenic plants, prokaryotic and eucaryotic expression systems. Methods for detoxification of grain, grain processing, silage, food crops and in animal feed and rumen microorganisms are also disclosed.	
XX	Key	Location/Qualifiers
XX	sig_peptide	/feat: a
XX	prod_1	/product: "barley alpha-amylase signal peptide"
XX	prod_2	/product: b
XX	prod_3	/product: "Fumonisin-esterase-K11APV0 fusion"
XX	misc_feature	/note: "extra lysine codon"
XX	wgizumma159-A1.	
XX	27 JAN 2000.	
XX	08-JUN-1999:	99WJ-US15454.
XX	15-JUN-1998:	99US-0092946.
XX	21-MAY-1999:	99US-0135491.
XX	(PION ) EIONEER HI BREU INT INC.	
XX	(CURA-) CURAGEN CORP.	
XX	Pavick JP, Gilliam JP, Madrox JR;	
XX	WEI, 2000 18:257-715.	
XX	P-PBDH; AAY58915.	
XX	New isolated polynucleotides, polypeptides useful for detection and degradation function of structurally related mycotoxin in processed grain or in silage	
XX	Example 13: Page 109-113; 154pp; English.	
XX	The present sequence is that of a polynucleotide encoding a fusion protein (see AAY58915) composed of the barley alpha amylase signal peptide, and a bacterial fumonisin esterase (BSV1) joined via a peptide linker to K11APV0 (see AAY58915), a truncated, but functional, amino polyol amine oxidase of Exophiala spinifera. The construct was designed for expression and secretion in plants of a hybrid Petunia having BSV1 and APDO activity. The invention provides APDO polynucleotides (see AA758383-87) and polypeptides (see AA758900-05) of P. spinifera and uninfectious derivatives. The polynucleotides are used to transform plant cells normally susceptible to Fusarium or other toxin-producing fungus infection. Also provided are methods for expressing APDO, optionally as a fusion protein with fumonisin esterase, in transgenic plants, prokaryotic and eucaryotic expression systems. Methods for detoxification of grain, grain processing, silage, food crops and in animal feed and rumen microorganisms are also disclosed.	







914 CCGACGCGTAAAGCGCGGCGATCGCGGCGCGCGACCTGCGCGCTTCGG 963  
 440 VALLHVAAGGCTGThVAlAlAsAsAsAsAsAsAsAsAsAsAsAsAsAs 346  
 964 GTCGCGATGCGAAATGCGCGAGCGAGCGAGCGCGCTTC...CTCGCGCG 1010  
 446 G...AsuAspThGlnAlaLysLeuGlnGlnAlaAla 357  
 1011 GCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1060  
 447 LpProAsuGlnLpProAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 373  
 1061 TCG 1101  
 448 GlnSerProGlnGlnGlnSerProGlnGlnGlnGlnGlnGlnGlnGlnGln 390  
 1102 ...GAGCG 1145  
 449 LpLpValAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 407  
 1146 CCGACATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1195  
 407 SNAAGGCTGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 423  
 1196 GCG 1245  
 424 LpGlnLpLeuPheProGlnSerGlnValGlnGlnGlnGlnGlnGlnGlnGln 440  
 1246 ...GAGCG 1283  
 440 VALLHVAAGGCTGThVAlAlAsAsAsAsAsAsAsAsAsAsAsAsAs 447  
 1284 GCG 1333  
 442 GlnAlaSerAlaLpLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 463  
 1334 CCGACG 1383  
 444 AlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 474  
 1384 GCG 1433  
 474 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 491  
 1434 GCG 1483  
 491 LysLysAlaLpLeuAlaLpValSerProAlaLpLpLeuAlaLpAlaLpAlaLp 507  
 1484 GCG 1527  
 508 AlaLpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 524  
 1528 GCG 1544  
 524 LpLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 541  
 1545 GCG 1594  
 541 AlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 557  
 1595 LpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1644  
 558 AlaLpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 574  
 1645 GCG 1694  
 574 LpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 591  
 1695 LpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1744  
 591 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 607  
 1746 GCG 1794

608 ValSerAlaLeuPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 624  
 1795 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 1844  
 624 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 641  
 1845 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 1894  
 641 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 657  
 1895 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 1944  
 658 LpLeuPheValLpPheSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 674  
 1945 CCG 1994  
 674 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 691  
 1995 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2044  
 691 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 707  
 2045 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2094  
 708 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 724  
 2095 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2144  
 724 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 741  
 2145 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2194  
 741 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 757  
 2195 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2244  
 758 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 774  
 2245 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2294  
 774 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 791  
 2295 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2344  
 791 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 807  
 2345 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2394  
 808 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 824  
 2395 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2444  
 824 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 841  
 2445 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2494  
 841 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 857  
 2495 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2544  
 858 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 874  
 2545 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2594  
 874 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 891  
 2595 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2644  
 891 LpLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLpAlaLp 907  
 2645 GATCGCGATGTTGTTGAAAGATTTCATTGCGAGCGCGCGCGCGCGCGCGCG 2694









[illegible]

[illegible]



[illegible]

11 New isolated polyphenol-oxidase polypeptides useful for detecting and  
p1 deterring fumonisin or structurally related mycotoxin in processed  
p1 grain or in silage -  
xx  
xx Example 10, Page 97-100; 154pp; English.  
xx  
xx  
xx The present sequence is that of the coding region of full-length  
cc amino polyol amine oxidase (APAO) cDNA of *Xophthalma splutifera*  
cc 214110. APAO is capable of degrading fumonisin. Its hydrolysis  
cc product AP1, and related mycotoxins. The invention provides APAO  
cc polyphenol-oxidase (see AA253281-87) and polypeptides (see AA556400-05)  
cc *E. splutifera* and *Rhizoctonia stolonialis* atrovitrosus. The polyphenol-oxidase are  
cc used to transform plant cells normally susceptible to fusion or  
cc other toxin-producing fungus infection. Transgenic plants can be  
cc regenerated from the transformed plant cells. Also provided are  
cc methods for expressing both APAO and a fumonisin-esterase in a  
cc transgenic plant, and for producing APAO enzyme in prokaryotic and  
cc non-plant eukaryotic systems. Transgenic plants capable of degrading  
cc fumonisin or of producing the degrading enzymes are provided.  
cc Methods for detoxification of grain, grain processing, silage, food  
cc crops and in animal feed and rumen microorganisms are also disclosed.  
cc APAO polypeptide is also useful as a selectable marker.  
xx  
xx Sequence 1808 BP; 424 A; 501 of 502 of 476 T; 0 other;

```

alignment_scores:
  Quality: 2392.50
  Ratio: 4.455
  Length: 633
Percent Similarity: 84.834
Percent Identity: 77.883

```

alignment\_block;

Alt. seg 1/2 L. M. 58402 From: 1 to: 1803

[illegible]









[illegible]









1325 GATGCAAAATAGGATATGAGTCGATTTGCCATGCCATGTCAAGGAACTT 1374  
767 ValProGlySerValHisLeuAsnThrProValAlaGluIleGlnIse 783  
1375 GTTCAGACCTCACTGCACCTTAAATATTCGTGCTGGAATTGCACAGTC 1424  
783 PAlaSerGlyCysThrValArgSerAlaSerGlyAlaValPheArgSerI 800  
1425 GGCAATCCGCGTGTACATACCATCGCGCTCCCGCGGTTCGGAAGCA 1474  
800 yslYsValValSerLeuProThrThrLeuThrProThrLeuThrPhe 816  
1475 AAAAGGTGGTGGTTGTTACCCGACACCTTGATCCACCTTGACATT 1524  
817 SerProProLeuProAlaGluLysGlnAlaLeuAlaGluAsnSerIle 833  
1525 TCACCACTCTTCCCGCTGAGAAAGCAACCATTTGGCGAAATTCATCT 1574  
833 uGlyTyrTyrSerLysIleValPheValTPAspLysProThrPArg 850  
1575 GGCTTACTAIAAGCAAGATAGTCTTGATGGAGACAGCCGTGGCGCG 1624  
850 IugInGlyPheSerGlyValLeuGlnSerCysAspProIleSerPhe 866  
1625 AACAAAGCTTCTCGGCGCTCTCCATCGACCTGTGACCCCATCTCAT 1674  
867 AlaArgAspThrSerIleAspValAspArgLutIlePheSerIleThrCysPh 883  
1675 GCCAAGCATACCAAGCATGACGTGATGACATGACATGCTCATTACTGTT 1724  
883 GMeValGlyAspProGlyArgLysTyrPheSerGlnIseLysGlnVal 900  
1725 CATGCTCCGAGACCGCGGAGCCGACACTCTCCCAACAGTTCAGGAGTAC 1774  
900 TrGlnLysSerValIlePheAspGluLeuArgAlaIleTyrGluAsnAlaGly 916  
1775 GACAAAGTGTGTGTGGAGCAACTCCGCGCAGCTACGAGAGACCGCGG 1824  
917 AlAGlnValProGluProAlaAsnValLeuGluIleGluTyrPheLysG 933  
1825 GTCGAGCTCCCAAGCGCGCCCAACGTCGCAAAATCGAGTGTGCAAGCA 1874  
933 uGlnTyrPheGlnGlyAlaProSerAlaValTyrGlyLeuAsnAspLeu 950  
1875 GCAATATTTCCAGAGAGCTCCGAGCCCGCTATGGGTGAAGCATCTCA 1924  
950 IeThrLeuGlySerAlaLeuArgThrProPheLysSerValHisPheVal 966  
1925 TCCACTGTGGTTGGCGCTCAGAACCGCTTCAGAGTGTTCATTTCGTT 1974  
967 GlyThrGluThrSerLeuValTPPLysGlyTyrMetGluAlaIleLeu 983  
1975 GCAACCGAGAGCTTACTTTGAAAGGCTATATGCAATGGCGCATACG 2024  
983 ySerGlyIleArgIValAlaAlaGluValValAlaSerLeuValProAla 1000  
2025 ATTCGTGCAACGAGGTGTGCAAGATGTGGCTAGGCTGTGTCAGCAG 2074  
1000 Ia 1000  
11  
2075 CA 2076





[illegible]

[illegible]

Quality:	5204.00	Length:	1000
Ratio:	5.204	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

```
alignment_block;
```

Accession No. 1/1 to: 45-09-352-168-24 From: 1 to: 3003

[illegible]

267	rtrhvalaiaaleuaspysrthlrhserileaspjileuenser	284
801	cacgtgtcaacgtcttcacactcaacacacgtatagacattttgaatttga	850
284	elafgagvalasplevalatthleketasnthrileugileugly	300
851	tgacaacagctcgatctcgacactctatgaacacgatcagcaactcgca	900
301	leuglyphegltufrrtrleuaspasvaltthalaalttarafereng	317
901	cttgagcttttgatgacaggttggaacgtaacgctgtgacgtttcga	950
317	uthralarothrthrcgiaspillealafgvalprovallevalg	334
951	aacgcttcgacacgactgatacatgtctgtgacgtgtctctgtcgga	1000
334	hrvalaiaaspaspjileuuephevalleuglyuasnaspthrgh	350
1001	cggggcccaacgacgacactctcttttttctcctcagaaatcgaaccca	1050
351	alatyrlleuglugaalileproasndiproaspheurygrinhrle	367
1051	gtaatvtrpamamssaaatpctgaatpaaatpcttttaccagactct	1100
367	aleuglyalalytrrlllecllysertrpodyllecylserproginasp	384
1101	ccttgagacataatgccattgatccccagagatgcattcctccacacac	1150
384	initlealaahtleugluthghvalmrphedhncsproseralale	400
1151	agattggccccattgagaccagaggaatgaatattcagtgcccttgccatc	1200
401	valaiaiaaspsealagaaiaaaglythleoseertrpafrrtyrry	417
1201	gtggcttcaggaactccgggaatcggggatccctcttgagcgactactcta	1250
417	rasnaalathrhecluasnuleuglucupheproglysercdhvaltyrh	434
1251	caatcgacacttttgcgaattctgggaatttttttttttttccgaactgac	1300
434	lsserserglvalaelymealvalphecllytrtyrprovalaiaaser	450
1301	acagctctcgaaatcgagatgagttgtttggcagcgmccgtgcgaagcgcg	1350
451	thrAlaleuglualaginthserlystymecolnolgalairpaiaa	467
1351	acggcttgaggaaacccaaatagaaataatctgaggtgtgtgtttgtttg	1400
467	aphealialysnprometastaglyproglittprplyscdhvalproasn	484
1401	cttttgccaaaaaacccatgatgagggcctgggcttggaacaaatggcgatg	1450
484	alalalaiaaueuglyserprocllyysalalileginalasfpvalser	500
1451	tccggcgcccttggtgcacacagcaaaagccaaacagttgacgtcttcca	1500
501	alathrileaspghargysalaleuhyrthrafertryrtrrpsule	517
1501	ggcaacaatagacaaacagatgctcttgacacggcttatattactgactt	1550
517	ucllythlialaiaaeragthrtheclytytylserctlytytylserg	534
1551	ggagcaaaatccgcgcgaagatctttggcgagagatgagggatgaaatg	1600
534	lygltytylserctlystasasavaiaaasfpvalvalvalaiaaiaa	550
1601	gcgcgaaggaacgaaacacaaatcttctcgaatgagaaatgagaaatgaa	1650
551	leuserglileugluthalalarqysvalalinalaiaaiaaileuser	567
1651	tttagcgcgcttgagaaacggcgaacgaaagctcagagagatgctgctg	17000
567	stcuvalleuglualaiaaetaspargvalaiglylysthrleuservalg	584

```

|||||
1701 CCGCGCTTCTGAGGAGATGATCGTGTAGGGGAAACATCTGAGGTGATC 1750
584 InsrGlyProGlyValGlyThrThrIleAspAspLeuValAlaAlaTrpPhe 600
|||||
1751 AATCGGCTGCGGCTAGCGACGACATCAAGCAACCTCGGCGCTGGCGGATC 1800
601 AspAspSerAsnGlnSerGluValSerArgLeuPheLeuArgPheHisIle 617
|||||
1801 AATGATGATGACCAACGCAAGATATCGAGATTCGTTAAAGATTTCATTTC 1850
617 GlnGlyIleLeuGlnArgThrThrPheAsnSerIleHisGlnAlaGlnA 634
|||||
1851 GGAAGCGAGCGTCTAGAGGAGGAGCTGAAATTCATATCAACACCAAG 1900
634 SpGlyThrThrThrAlaIleProGlyGlyAspSerLeuLeuSerGlnGln 650
|||||
1901 ACGGAGACCGCATACAGCTGCTTATGCTGATTCCTTCTGAGGAGGAG 1950
651 ValAlaSerAlaLeuAlaGlnLeuLeuProValTrpSerGlnIleLeuIle 667
|||||
1951 GTTCGAAATGCACTTCCGAACTCTCTCCGATATGGTTCATATCAATCA 2000
667 GlnGlySerLeuThrAspLeuValAspSerProGlnAlaIleAspLeuVal 684
|||||
2001 AGGCAATAGCTTCTAAGACTTCAAGCTGAGCTCTCAAGCTCAAGTGGCT 2050
684 SpSerValSerPheAlaHisIleTrpGlySerIleLeuValAsnLeuProAla 700
|||||
2051 ACATGAGACCTTCTGCGACTACTGTCAGCAACCAATCAACCTTCCTGCT 2100
701 ValLeuGlyValAlaAsnGlnIleThrArgAlaLeuLeuGlyValGlnAla 717
|||||
2101 GTTCTGCGGCTGATCAAGCAAGATCAACGCTGCTGTCTGTCTGTGAGCT 2150
717 GlnSerIleLeuSerMetIlePheLeuThrAspTrpTrpGlySerAlaIle 734
|||||
2151 CACGCAATACGCAATGCTTCTTCACGCGCATACATCAAGTGGCGAGCG 2200
734 LysLeuSerAsnIlePheSerAspGlySerGlyGlyGlyTrpMetArg 750
|||||
2201 GTCTCATTAATATTCTCTCGCAACCAAGCAAGCTGCTGAGTATATGCA 2250
751 CysGlyThrGlyMetIleSerIleGlyHisAlaMetSerGlySerGlyGly 767
|||||
2251 TGTAAAGACGATGAGTCAATTCGATTCGATTCGATTCGATTCGATTCG 2300
767 TrpGlySerValHisIleAsnThrProValAlaGlnIleGlnIleSerA 784
|||||
2301 TTTAGGCTTAGCTGAGCTCAAGCTCAAGCTGAGCTGAGCTGAGCTGAG 2350
784 LysArgIleGlyThrValAlaSerAlaSerGlyValAlaValPheAspSer 800
|||||
2351 CATCGCGCGCTGATAGTAAAGACGCTTCTGCGCGCGCTGCTGCGCAAA 2400
801 LysValValAlaSerLeuProThrIleLeuTrpProIleLeuThrPheS 817
|||||
2401 AAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2450
817 TrpProIleProAlaGlnIleGlySerAlaLeuAlaGlnIleAsnSer 834
|||||
2451 ACCAGCGCTGCGGCTGAGGAGCAAGCTTCTGCGCAAAATTCATCTGCG 2500
834 LysTrpSerIleValPheValTrpAspGlyProIleTrpArgGln 850
|||||
2501 GCTTCTATATATATATATATATATATATATATATATATATATATAT 2550
851 GlnGlyPheSerGlyValLeuGlnIleSerSerGlyAspProIleAspPhe 867
|||||
2551 CAAGGCTTCTCGAGGCTGCTGCAATGAGCTTCTGAGCTCAATCTTCTG 2600
867 LysArgSerIleThrAspValSerArgGlyGlySerIleThrArgPhe 884
|||||

```

```

2601 CAGACATACCAACATCGACGCTGATCGACATAGTGCATTACCTGTTTCA 2650
884 eValGlyAspProGlyIleArgLysTrpSerGlnGlnIleSerGlySerValAla 900
|||||
2651 TGGTGGAGACCGCGGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGT 2700
901 GlnLysSerValTrpAspGlnLeuArgAlaAlaIleArgIleGlnAsnAlaGly 917
|||||
2701 TAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2750
917 GlnValProGlnIleProAlaAsnValLeuGlnIleLeuTrpSerLysGln 934
|||||
2751 CCAAGTCTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2800
934 IsoleuGlySerAlaLeuArgThrProPheLysSerValHisPheValG 967
|||||
2801 AGTATTTCAGAGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 2850
967 TrpLeuThrSerLeuValTrpGlyTrpMetGlnGlyValAlaIleArg 984
|||||
2851 ACAGTGGCTTCTGCGCTGAGAACTTCTGCGCTGAGAACTTCTGCGCT 2900
2901 AACGCAAGAGCTTCTTAGTGTGCAAGCGGTATATGCAAGCGGTATAT 2950
984 eGlyGlnArgGlyAlaAlaGlnIleValValAlaSerIleValProAlaAla 1000
|||||
2951 CAGGTCAACGAGAGTCTGCAAGAGTGTGAGTGTGAGTGTGAGTGTG 3000

```

seq\_name: /cqn2\_6/ncdata/2/ins/chr25m.seq.us-09-652-152-28

seq\_documentation\_block:

```

: sequence 28, Application US/0932159A
: Patent No. 6211434
: GENERAL INFORMATION:
: APPLICANT: Duqvick, Jonathan P.
: APPLICANT: Gilliam, Jacob T.
: TITLE OF INVENTION: Amino Polycl Amine Oxidase
: TITLE OF INVENTION: Polynucleotides and Related Polypeptides and Methods of use
: FILE REFERENCE: 1134
: CURRENT FILING DATE: 1999-07-12
: EARLIER APPLICATION NUMBER: US/09/352,159A
: EARLIER FILING DATE: 1998-07-25
: EARLIER APPLICATION NUMBER: 60/135,391
: EARLIER FILING DATE: 1999-05-21
: NUMBER OF SEQ. ID NOS: 46
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ. ID NO 28
: LENGTH: 3618
: TYPE: DNA
: ORGANISM: Unknown
: FEATURE:
: OTHER INFORMATION: gsc:espl.sp.k:trapo.3618.1-687.98T
: OTHER INFORMATION: polylinker.599.2590.espl.mat.2191.2226.spacer
: OTHER INFORMATION: 2227-3615. K:TRAPO, 3616-4618, Stop codon. For
: OTHER INFORMATION: bacterial expression.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(3615)
: NAME/KEY: misc_feature
: LOCATION: (1)...(687)
: OTHER INFORMATION: gsc + polylinker
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: (688)...(2190)
: OTHER INFORMATION: espl.mat
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (2191)...(2226)
: OTHER INFORMATION: spacer sequence

```

FEATURE:  
 NAME/PROT: misc\_feature  
 LOCATION: (222)...(3615)  
 OTHER INFORMATION: K1:R1AP0  
 FEATURE:  
 NAME/PROT: misc\_feature  
 LOCATION: (222)...(2229)  
 OTHER INFORMATION: Extra lysine  
 US-09-658-159-28

alignment\_scores:  
 Quality: 5099.00 Length: 1016  
 Ratio: 5.130 Gaps: 2  
 Percent Similarity: 97.835 Percent Identity: 96.752

alignment\_block:  
 US-09-658 835c-25 x US-09-352-159-28

Align seq 171 to US-09-352-159-28 from 1 to 3618

```

3 AsnLysHisHisGlySerLeuSerLeuPheLeuVal... LeuLeuGlyLe 17
   |||||
568 GATTAAGTACTTCAATCAATCAAGAAATATAGCAATGAGCTTTGACAGCTG 617
17 GSEALASerLeuAlaSerGly... 24
   |||||
618 GGAAGGACGCTTTGCTGTGACCTATCCCAAAATCGATCGCTTC 667
25 ..... AlProThrValHisHisPheAlaGlyMet 34
   |||||
668 GCGCTGATCCCGGAAATTCGCTCTCTCTCAAGATTGATGCTGGGATG 717
45 ValValHisThrThrThrThrValProGlyThrThrAlaThrValSerG 51
   |||||
718 CTCTGCGGACGACTATCACTATGTCGCGACCACTGCGACCGCTCAGCA 767
51 UpheLeuGlyValProPheAlaHisProThrValPheAlaProPro 68
   |||||
768 GTTCTTGGGCGTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 817
68 hTrpPheValProPheSerThrProLeuGlnAlaThrAlaTyrGlyPro 84
   |||||
818 CTGTCGCGGCTGCTTGGTTCACAGCGCTTGTCAAGGCACTGCATATGTC 867
85 AlaGlySerGlnGlnProPheAsnTyrProGlnLeuArgGlnLeuPhe 101
   |||||
868 GCAATGCTCTCAACCAATTCAATTACCCGGAAPACTCCGAGATTCGAT 917
101 LAlaThrPheAsnThrProProPheSerAlaGlyGlySerGlnAspGly 118
   |||||
918 GGCGCTTTAAATATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 967
118 GAspLeuAsnLeuTyrValTyrGlyThrGlnLeuSerThrAsnGlyAla 134
   |||||
968 TGAACCTCAACATCTAGCTCCAGAGCACTGAGAACACAAACAAAGCC 1017
135 MetValTyrPheTyrGlyValAlaLeuGlyTyrGlyTyrPheAsnSer 151
   |||||
1018 ATGCTTTCTATATACCTCGAGCTCTGCAATATGCTTGAATTCATTC 1067
151 SLeuTyrAspTyrAlaSerProPheAlaHisAsnGlnAspValHisAla 168
   |||||
1068 CATTACAGAGGCTCTATTTCTGACCAATAGCATGCTCATGCGCGTGA 1117
168 hTrpLeuAsnTyrAlaGlyTyrGlyThrGlnLeuSerThrAsnGlyAla 184
   |||||
1118 CCATCAACATCAAGAGCAACATCTGAGGCTTCTGCTGCTGCTGCTGCT 1167
185 ProLeuThrGlnArgAsnLeuGlyPheLeuAspGlnIleArgPheAlaLeu 201
   |||||
1168 GCAATATACAGAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1217

```

```

201 PTrPheAlaGlnArgAsnLeuAlaAlaPheGlyGlyAspProArgGlyVal 218
   |||||
1218 TGGGCTCAACAGGAGAAACAGCAACGCTTTGCGCTGATGCTCAAGGCTCA 1267
218 hTrpPheGlyGlnSerAlaGlyGlyValArgSerValAspValLeuLeuThr 234
   |||||
1268 CAATATTTGGGCTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317
235 SerMetProHisAsnProPheArgValAlaAlaIlePheGlySerGlyVal 251
   |||||
1318 TCTATGCCACACAAACCCACCTCTCGAGGACCAATCAATGAGGCTCGGCT 1367
251 LAlaSerTyrAsnProPheGlyGlyAspLeuSerGlyAspProPheAsnThr 268
   |||||
1468 AGAAGAGTGGATCTTACCTCTTATGACACAGATCGAGCACTGAGCT 1517
301 GcLysPheGlyTyrThrLeuAspAsnValThrAlaValTyrArgSerGln 318
   |||||
1518 TGGGTTGAGTACAGCTTGGACACAGTAAAGGCTGTGACCTTCTGAA 1567
318 hTrpAlaArgThrThrGlyAspLeuAlaGlyValProValLeuValGlyThr 334
   |||||
1568 CGGCTGAGACAGTATGAGACATGCTCGTACGCTGCTGCTGCTGCTGCT 1617
335 ValAlaAsnAspGlyLeuLeuPheValLeuGlyGlyAsnAspThrGlnAl 351
   |||||
1618 GTGGGCAACAGAGGAGCTCTCTTCTCTCGCCGCAAAATGACCCAGAG 1667
351 aTyrLeuGlnGlnAlaIleProAsnGlnPheAspLeuTyrGlnThrLeu 368
   |||||
1668 ATATCTTATGAGGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1717
368 GcLysAlaGlyThrProLeuGlySerPheGlyTyrGlySerProLeuAsn 384
   |||||
1718 TTGAGCATATCCCATGATGATCCCGAGGATCGGATCGGCTCAAGATG 1767
385 LLeuAlaAlaIleGlnThrGlnValAlaGlyPheGlnGlyProSerAlaLeu 401
   |||||
1768 ATGCGCGCATGAGAGCGAGGATGAGATTCAGTGTCTGCTGCTGCTGCT 1817
401 LAlaGlnAsnSerAlaAsnAlaGlyTyrProSerThrPheArgTyrTyrP 418
   |||||
1818 GGCTCAAGAGCTCCCGAATGCGGATGCGGATGCGGATGCGGATGCGGAT 1867
418 SAlaIlePheGlnAsnLeuGlnLeuPheProGlySerGlnValTyrHis 434
   |||||
1868 ATGCGACCTTTGAGAACTCGAGCTTTCCCGGCTGCAAGGCTGACAG 1917
435 SerSerGlyValGlyMetValHisPheGlyThrTyrPheValAlaSerAla 451
   |||||
1918 AGCTCTGAAGTGGAGATGCTTGTGAGACGTATCTGCTGCTGCTGCTGCT 1967
451 PAlaLeuGlnAlaGlnIleSerLysTyrMetGlnGlyAlaIlePheAla 468
   |||||
1968 GCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2017
468 hLeuAlaAsnProMetAsnGlyTyrGlyTyrPheGlnValProAsnVal 484
   |||||
2018 TTCCCAAAAACCCCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2067
485 AlAlaLeuGlySerProGlyTyrAlaIleGlnValAspValSerProAl 501
   |||||
2068 GCGGCTGCTTGGCTGATCAAGCAAGGATCCAGGTGAGCTCTCTGCAAG 2117
501 aThrIleAspGlnArgGlyAlaLeuTyrThrArgGlyTyrThrGlnLeu 518

```



```

OTHER INFORMATION: 2227-3615, K12ALVA, extra lysine 466-468, stop
OTHER INFORMATION: codon for bacterial expression
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(3615)
NAME/KEY: misc_feature
LOCATION: (1)...(466)
OTHER INFORMATION: qast + polylinker
FEATURE:
NAME/KEY: misc_feature
LOCATION: (688)...(2190)
OTHER INFORMATION: espi mat
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2151)...(2226)
OTHER INFORMATION: spacer sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2227)...(3615)
OTHER INFORMATION: K12APAO
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2227)...(2229)
OTHER INFORMATION: Extra lysine
US-09-652-168-28

alignment_scores:
Quality: 5099.00 Length: 1016
Ratio: 5.130 Gaps: 2
Percent Identity: 97.835 Percent Identity: 96.752

alignment_block:
US-09-658-835c-25 x US-09-652-168-28
Align: 501 1/1 100 98-09-352-168-28 from: 1 to: 3618
3 ASDLSHLSLSerLeuSerLeuPheLeuVal.....LeuLeuGlyLe 17
568 GATACACTACTTCAATCCAGCAAGATATAGCATGCCCTTTCAGAGGCG 617
17 uSerAlIserLeuAlaSerGly..... 24
618 GCAGGCCAGGTTGGTGGTGGGACCAATCCCAAAATCGATGGTTC 667
25 .....AlaProthrValIysI]easPalaglymet 34
608 CCGGCGGATCCCGGAAATTCATCTGCTTCAATTTATTTATTTGATG 717
35 ValIValGlyThrThrThrValIProGlyThrThrValIserGly 51
718 GTGGTGGGACGACTACTACCTGGGAGGAGAAATGAGAAATTTASGA 767
51 uPheLeuGlyValIProPheAlaIAserProthrArqPheAlaPro 68
768 GTTCTGGGAGTTCCTTTCGAGGATTCATGATTTAGGCTCTCA 817
68 hArqProValIProI]SerThrProLeuGlnAlaThrAlaTyrGlyPro 84
818 GTGGTGGGAGTTCCTTTCGAGGATTTGATGATTTAGGCTCTCA 867
85 AlaGlyProGlyGlnI]PheAsnTyrProGlyGlyLeuArqGlyI]e 101
868 GCATGCCGCTCAACATTCATTAATTAATTAATTAATTAATTAATTA 917
101 lAlaI]PheAsnThrProProSerAlaGlyI]uSerGlyAspCysLe 118
918 GGTCTGGTCAATACACGCGCGCGGATTCATGATTTAGGCTCTCA 967
118 euAsnLeuAsnI]eTyrValIProGlyThrGlyuAsnThrAsnI]yAl 134
968 TGAACCTCAACATCTACGTGGAGGAACTGAGAACACAAATCAATCC 1017

135 MetValITrTLeuTyrGlyValAlaLeuGlyThrGlyTyrAsnSerPhe 151
1018 ATGGTTGGATATATGCTGGAGCGCTGGAAATAGTTGCAATTCATTC 1067
151 sleuTyrAspGlyAlaSerPheAlaAlaAsnGlnAspValI]leAlaVal 168
1068 CTTTTCATGATGATTCATGATTCATGATTCATGATTCATGATTC 1117
168 hrlIeAsnTyrArqThrAsnI]eLeuGlyPheProAlaAlaProGln 184
1118 CCATCAACATCAACATCAACATTCATGATTCATGATTCATGATTC 1167
185 ProI]eThrGlnAlaGlnLeuGlyPheLeuAspGlyArqPheAlaCys 201
1168 CCAATATATATATATATATATATATATATATATATATATATAT 1217
201 pI]I]ValGlnAlaGlnI]leAlaAlaPheGlyGlyAspProAlaGlyVal 218
1218 TGGGAGACCAAGGAGATTCATGATTCATGATTCATGATTCATG 1267
218 hrlI]PheGlyGlnI]SerAlaGlyI]eTyrSerValI]AspValI]leuLeuThr 234
1268 CAAATATTTGGGACAGTTCGCGGAGGAGAGAGTTCGACGCTCTTC 1317
235 SerMetProI]AsnProProPheArqAlaAlaI]eMetGlnSerGly 251
1318 TCAATGCCACCAACCCACCTTCGACGACATATCATGATGATGAT 1367
251 lAlaAsnTyrAsnProI]eTyrGlyAspLeuSerGlyProI]AsnThr 268
1368 GCGTACTACACTTCGCCAGGAGATTTGTCGCAACCTTGAACACCA 1417
268 hValI]GlnAlaI]eAsnCysThrThrSerI]eAspI]eLeuSerCysMet 284
1418 CTTTTCAGGCTCAACCTTACACCATGACCATGACATCTGATGATG 1467
285 ArqArqValAspLeuAlaThrI]eMetAsnThrI]eLeuI]eLeuGly 301
1468 ACAAACATTCATGATTCATGATTCATGATTCATGATTCATGAT 1517
301 uGlyPheGlyThrThrLeuAspAsnValI]ThrAlaValI]TyrArqSer 318
1518 TGGCTTTCATATACGTTTGGACACGATACGCTGATGATGATGAT 1567
318 hrlAlaArqThrThrGlyAspI]eAlaArqValI]ProValI]leuValI]GlyThr 334
1568 CGGCTGCCACACGCTGATGATTCATGATTCATGATTCATGATTC 1617
335 ValAlaAsnAspGlyLeuLeuPheValI]eGlyGlnAsnAspThrGln 351
1618 GTGGCCACACGAGGACTTCCTTCTGCTCGGAGATTCACACCCAG 1667
351 aTyrLeuGlyGlnAlaI]eProAsnGlnProAspI]eThrGlnThrLeu 368
1668 AATTCCTGAGGAGATTCATGATTCATGATTCATGATTCATGAT 1717
368 euGlyValI]TyrProI]eGlySerProGlyI]eGlySerProGlnAspGln 384
1718 TTGGAGCATATCCATTGATCCCGAGGATTCGATGATTCGATGAT 1767
385 l]leAlaAlaI]eGlyThrGlnValAlaArqPheGlnCysProSerAlaI]leVal 401
1768 ATGGCCGCTTGGAGCGAGGATTCATGATTCATGATTCATGATTC 1817
401 lAlaGlnAspSerAlaAsnValI]eTyrSerTyrArqTyrTyrTyr 418
1818 GGTCTGGAGTTCGCGGATTCATGATTCATGATTCATGATTCATG 1867
418 snAlaThrPheGlyAsnLeuGlyLeuPheProGlyI]eSerGlyValI]TyrHis 434
1868 ATGCACCTTTGACATCTGACCTTTTCCGAGGATTCGAGAGTTCAC 1917
435 SerSerGlyValI]GlyMetValI]PheGlyThrTyrProValAlaSerAlaThr 451

```



```

|||||
1918 AGCTGTGAAGTGGGATGGGCTTTTGGGAGTATGTTGGCCAGGCTGGAC 1967
451 FALGPHGLHLAGLHDSSTSTYMETGGLVALATPHALALAP 498
1968 GCGCTGGAGGGGTAGGCGAGGAGAAATATCGAAGGTCGCTGGCGGCGCT 2017
498 HVALHYSASDPTMETASDLYTHGDIYPLYSGLVALPPLASVAL 484
2018 TGGCAAAAGAAAGCATAAATAGGAGTGGTGGAAACAGAGTGGCAATGGC 2067
485 ALALALGGLYSCTPROLYLSALALGGLVALASVALSPTAL 501
2068 GCGGGGCTTGGCTTACCAAGCAAAAGCAACAGGCTTGGAGTGGTGGAC 2117
501 ALHLLASPTLALGYSALAGGLYTHFARGLYTHFARGLYTHFARGLY 518
2118 GACCAATACAAATATGCTGCTGGLAATAGCGTATATATATGAGTTGG 2167
518 LYTHLLALAPPRATThHDSSTLYGGLYSCTLYSCTLYSCTLYS 544
2168 GAGAAAGCGGGGAGAGCAATTTGGGAGAGAGAGAGAGAGAGAGAGAG 2217
545 GLYGYSSTYSASPASVALALASVALVALVALVALVALVALVAL 551
2218 GAGASVALASVALAGAAAGAGCTTGGAGAGAGAGAGAGAGAGAGAG 2267
551 GSETGLYGLUHLFARALATVALVALALALAGLYLGLYLSCTLYS 568
2268 GAGGAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2317
568 GVALVALGGLALALASPARVALGLYGLYLSCTLYSCTLYSCTLYS 584
2318 TGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2367
585 SETGLYGLYGLYATHTHTHTHGLASASPLGLYALALATPLGLAS 601
2468 TGGGGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2417
601 GASPSTASNDLSTASGLYVALSTASPLGLYVALPPLASPLGLY 618
2418 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2467
618 TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 634
2468 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2517
635 GLYPHTHTHTHTHGLASPLGLYASPSTGLYGLYGLYGLYGLY 651
2518 GGTAAACCAATTAACATCTGCTTATGGTCACTGGTGGAGAGAGAG 2567
651 LALASGLALALAGLGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 668
2568 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2617
668 LHLISSTLEGLASPLGLYASPLASPTGLYGLYGLYGLYGLYGLY 684
2618 AGCATAGCGCTTCAATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2667
685 SETVALSPTGLALALHLSLYTGLYSCTLYSGLYGLYGLYGLYGLY 701
2668 AHTGGAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2717
701 LGLGLYVALALASNDLSTHTHTHTHTHTHTHTHTHTHTHTHTHT 718
2718 TGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGG 2767
718 LSGDLHSETMETGLPHGLHGLASPTLYTHGLYSCTLYSCTLYSCTLY 734
2768 AGGAAATAGAGATGTTTHTGCAAGAGAGAGAGAGAGAGAGAGAGAG 2817
735 LGLSSTASGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 751
|||||

```

```

2818 GTAGTAAATATTTTCTGGACAAAGAGAGAGAGAGAGAGAGAGAG 2867
751 SLVSTHGLYMETGLISCTHTGYSHTLSALAMETSEGLYGLYGLYVALP 768
2868 CAAACAGAGATATGTAAGAGATTTGGCATATGCTATGTAAGAGAGAG 2917
768 TGLYSERVALHLSLEASDPTHTPROVALALAGLGLYGLYGLYGLY 784
2918 GAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2967
785 SETGLYGLYTHRYVALASGLASGLYVALVALPPLASGLYGLYGLY 801
2968 TGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGG 3017
801 SVALVALSPTLEUPPTHTHTHTHTHTHTHTHTHTHTHTHTHTHT 818
3018 GGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3067
818 TGLPGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 834
3068 GAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3117
835 TYTYSCTGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 851
3118 TACTATAGCAGAGATAGCTTCTGATGGAGAGAGAGAGAGAGAGAG 3167
851 NGLYPHESEGLYVALLEUGLINSERSETCYASPTLLESETPHEALAA 868
3168 AGAGTGTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3217
868 TGAAPTHTSERHTASPTVALASPARGLYHTPSEHTHTHTHTHTHT 884
3218 GAGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3267
885 VALGLYASPTGLYATHTHTHTHTHTHTHTHTHTHTHTHTHTHT 901
3268 GTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3317
901 DLYSSEVALTTPASGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 918
3318 AAGATGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3367
918 LVALPGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 934
3368 AAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3417
935 TYTPHGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 951
3418 TATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3467
951 TLEUGLYSERVALLEUGAGTHTHTHTHTHTHTHTHTHTHTHTHTHT 968
3468 AGTGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGG 3517
968 HGLYHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHT 984
3518 GCGAGAGAGATTTATGTTTGAAGAGAGATATGTAAGAGAGAGAGAG 3567
985 GLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 1000
3568 GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3615

```

seq\_name: /cgn2\_57/ptodaly\_47/lna/58\_09MB.seq-us-09-658-835c-25

# seq\_documentation\_block:

Sequence 26, Application US/09352159A  
Patent No. 6211434

## GENERAL INFORMATION:

APPLICANT: Duwick, Jonathan P.  
APPLICANT: Gilliam, Jacob T.

APPLICANT: Maddox, Joyce R.  
TITLE OF INVENTION: Amino Polypeptide and Related Polypeptides and Methods of Use



1061 ITGGCAACCAAGCC.....GGCGCGTGGCGCGCTGCTATCCCTC 1101  
 374 GlySerProGlyIleGlySerProGlnAspGlnIleAlaIleGluThr 390  
 1102 .....GACGGCCGAGGACGCGCCCAAGAAATGCTGGCGCGCTCTCGG 1145  
 390 fglValAlaGpPheGlnGlySerProSerAlaIleValAlaGlnAspSerAla 407  
 1146 CGAGAAATCAATTCATTCGGCGGCTCTGGCGCTTCTCGCAAGCCCTTGTC 1195  
 407 SPARGIlyIleProSerTyrPheTyrTyrTyrAsnAlaThrThrGlnAsn 423  
 1196 GCGAGAGAGGCGCGCGCTGTGCGTATCAATTCAGCGTATATCCGACGGT 1245  
 124 LeuGlnIleThrGlnGlySerGlnValTyrIleSerSerGlnValGly 440  
 1246 .....CGAAGAGAGGCGCGCTACCGCAAGCGAAGCGAATTCGCA 1283  
 440 ValPheGlyIleTyrTyr.....ProVal 447  
 1284 GGTITTCGGGTGTTCAGCTGAGCGAGATGGCTCTGTCGATGGTGGC 1333  
 447 AlaSerAlaThrAlaGlnGlnIleThrSerIleTyrMetGlnGly 463  
 1334 CGGAGCGCGCGCGCGCGCGAGCGCTGGCGCTGGCGCACTGATGCTCTC 1383  
 464 AlaThrAlaAlaPheValAlaLysAsn.....ProMetAs 474  
 1384 GCTTATATTCGCTTCTTATATATATGCTATCTGCTGCGGCGCGCTTAC 1433  
 474 nGlyProGlyIlePheGluValProAsnValAlaAlaLeuGlySerProG 491  
 1434 CTGGCTGGCTATTCAGCGCGAGCGAGCGACATGACATTCGGTCCGAGC 1483  
 491 LysAlaIleGlnValAspValSerPheAlaThrIleAspGlnIleArg 507  
 1484 GTTAT 1527  
 508 AluIleTyrThrArgTyrTyrThrIleGlnGlyIleIleAlaIleProArg 524  
 1528 GCG.....GATGGCGCCCAAGCG 1544  
 524 PheGlnGlyIleGlySerGlyIleGlySerGlyIleGlySerLysAspAsp 541  
 1545 GAG 1594  
 541 AlaIleAspValValValIleValIleGlyIleSerGlyIleGlnIleAla 557  
 1595 TTGGGAG 1644  
 558 ArgIleValIleAlaIleAlaIleGlnAspLysLeuValIleGlnAlaMetAs 574  
 1645 CGGAAAGTCGAGCGCGCGCGCTCTGGCGCTGCTGGCGCGAGCGATGGA 1694  
 574 PheValGlnGlyIleGlyThrLeuSerValIleSerGlyIleProGlyArgThr 591  
 1695 TGTGTATGGGAGAAACGTCTGAGCTACCAATCGCGTCCGCGCAGACCA 1744  
 591 IleIleAsnAspLeuIleAlaIleAlaIleAsnAspSerAsnIleSerGln 607  
 1745 GTATCAAGCGAGCTCGCGCTGGCTGATCAATGACAGAACCCAAAGCGCA 1794  
 608 ValSerArgIleuPheGlnArgPheIleGlnGlyIleGlnIleArgThr 624  
 1795 GTATCCAGATTCCTTCAAGAGATTCATTTGGAGCGCGAGCTCCAGAGAG 1844  
 624 ThrGlnIleAsnSerIleIleGlnAlaIleAspGlyThrThrThrIleArg 641  
 1845 GAGTGAAGAAATCAATTCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1894  
 641 nGlyIleValSerProGlnGlyIleValIleSerAlaLeuAlaGln 657  
 1895 GTTATGTGACCTGCTGGTGAATGAGAGCTTTCAGTGCACCTTGGGGA 1944

658 LeuLeuProValTyrSerGlnLeuIleGlnGlnIleSerLeuGlnAspLe 674  
 1945 GTCT 1994  
 674 nGlyIleSerProGlnAlaIleValArgLeuAspSerValSerPheAlaIle 691  
 1995 CAAG 2044  
 691 LysGlnGlyIleGlnLeuAsnLeuProAlaValIleGlyValAlaAsnGln 707  
 2045 ATGTGAG 2094  
 708 IleThrArgAlaLeuLeuGlyValGlnIleIleGlnIleSerMetLeuPhe 724  
 2095 ATTCAG 2144  
 724 LeuThrAspTyrIleLysSerAlaThrGlyLeuSerSerIlePheSerA 741  
 2145 TCTACCGAGCTACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2194  
 741 sPLysLysAspGlyGlyIleTyrMetArgCysLysThrGlyMetGlnSer 757  
 2195 ATTCAG 2244  
 758 IleCysHisAlaMetSerLysGlnLeuValProGlySerValHisLeuAs 774  
 2245 ATTTCCATCCCATGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2294  
 774 nThrProValAlaGlnIleGlnGlnSerAlaSerGlyCysThrValArgS 791  
 2295 CACCGCGCGCTCTCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2344  
 791 eAlaIleSerGlyAlaValPheArgSerLysLysValValSerLeuPro 807  
 2345 GCGCT 2394  
 808 ThrThrIleTyrProThrLeuThrPheSerProProLeuProAlaGly 824  
 2395 ATCACT 2444  
 824 sGlnAlaLeuAlaIleAsnSerIleLeuGlyTyrTyrSerLysIleValP 841  
 2445 GCAAGCTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2494  
 841 IleValTyrAspLysProTyrPheArgGlnGlnGlyIlePheSerGlyValLeu 857  
 2495 TGTATGTAG 2544  
 858 GlnSerSerCysAspProIleSerPheAlaArgAspThrSerIleAspVa 874  
 2545 CAATTCAG 2594  
 874 LAspArgGlnTyrSerIleThrCysPheMetValGlyAspProGlyArgL 891  
 2595 GTATTCAG 2644  
 891 LysTyrSerGlnGlnSerLysGlnValArgGlnLysSerValTyrAspGln 907  
 2645 ATCT 2694  
 908 LeuArgAlaAlaTyrGlnAsnAlaGlyAlaGlnValProGlnProAlaAs 924  
 2695 CTCGCGCAGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2744  
 924 nValLeuGlnIleGlnTyrPheSerLysGlnGlnTyrPheGlnGlyValArg 941  
 2745 GTTCT 2794  
 941 eAlaValTyrGlyLeuAsnAspLeuIleThrLeuGlnGlySerAlaLeuArg 957  
 2795 GCGCGCTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2844













49 1serAluPhenLeuGlyValAlaAspAlaIleAspProThr.....A 63  
747 CATGAGCTTCCGGAAATACCAATAGTAAAGGCGGCGGCGGCTAC 796  
63 rApheAlaProFroThiArgProValProIserThrProLeuAla 79  
747 CTGGAGAGCGCCCGAACACACCGCGCGCGCGCGCGCGCGCGCGCG 846  
80 ThrAlaIrrProValLeuGlyTrogIleGlyIleAsnValProLeuIle 96  
847 ACCCAATATGGCTGGAGTAC..... 867  
96 uArgGluIleThrMetAlaIrrPheAsnThrProProSerAlaGly 113  
868 .....TTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 910  
113 LuSerGluAspGlyLeuAsnLeuAsnIleTyrValPro...GlyThrIle 128  
911 TCAATGAGGACATCTCTCTTAACTCAACCAATCCGCGCGCGCGCGCG 960  
129 AsnThrAsnTyrAlaValMetValIrrProIrrGlyTrogAlaLeuGly 145  
961 CCG 1010  
145 rGlyIrrAsnSerPheHisLeuTyrAspGlyAlaSerPheAlaAsnG 162  
1011 CGGACCG 1057  
162 LuAspValIrrAlaValIrrProAsnTyrArgIleAsnLeuGlyIrr 178  
158 AGGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1107  
179 rProAlaIrrProGluLeu.....ProIle...ThrGluArgAs 190  
1108 TTCCG 1157  
190 rLeuGlyIrrPheLeuAspGlyIrrGlyIrrPheIleAsnAspIrrValIrrArgAsnI 207  
1158 CTAGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1207  
207 LeuAlaValIrrAspGlyIrrAspProValGlyValIrrIrrGluGlyIrrSer 223  
1208 CGCGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1257  
224 AlrGlyIrrGlyIrrAspSerValAspValIrrLeuIrrSerMetProIrrAsnThr 240  
1259 CGCGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1307  
240 rProPheAlaAlaAlaIrrMetGlyIrrGlyValAlaAsnTyrAsnThrP 257  
1308 TCTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1355  
257 rLeuGlyIrrAspLeuSerGlyIrrProIrrPheIrrIrrValGluAlaIrrAsn 273  
1336 .....CCAGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1400  
274 rGlyIrrThrSerIleAspIrr...LeuSerGlyMetArgArgValAspLe 289  
1381 TCG 1400  
289 uAlaIrrLeuMetAsnThrIrrGluGlyIrrLeuGlyIrrPheGlu... 304  
1431 AGCGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1460  
305 .....TyrThrLeuAspAsnValIrrThrAlaVal 313  
1481 TGGCGAGCGCGCGCTGGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCTGGCG 1530  
314 TyrArgSerGlyIrrAlaIrrGlyIrrThrGlyIrrAspIrrAlaValIrrPro 330  
1531 CAAATCGACAGAGCGGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGCG 1580

330 IleValGlyIrrValAlaAsnAspGlyIrrLeuPheValLeuGlyIrr 346  
1581 CTGATGCGGAACCAAGAGCGAGGAAAGCGCGCGCGCGCGCGCGCGCG 1627  
347 .....AsnSerThrAlaIrrThrLeuGluAlaIle 357  
1628 CGCGCGTGAAGATATTAAGCGAGCAAAAGCGCGCGCGCGCGCGCTTT 1677  
358 rProAsnGluProAspIrrLeuGlyIrrThrLeuLeuGlyAlaIrrProIrrG 374  
1678 GCG 1717  
374 ySerProGlyIrrGlySerThrGluAsnGluIrrAlaIrrIrrGluThrG 391  
1717 .....CAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1765  
391 lValAlaPheGluIrrProSerAlaIrrValAlaGluAsnSerArgAsn 407  
1763 ACAATGATCTTAATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1813  
408 ArgGlyIrrProSerThrAlaIrrTyrTyrAsnAlaThrPheGluAsnLe 424  
1813 CAGCG 1861  
424 uGluLeuPheProGlySerGluValIrrHisSerSerGluValGlyMet 441  
1861 .....GAAATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1900  
441 aPheGlyIrrThrTyr.....ProVal 447  
1901 TTTTCG 1950  
448 AlaSerAlaIrrAlaIrrLeuGluAlaGluIrrThrSerLysTyrMetGluVal 464  
1951 GAGCG 2000  
464 rArgPheAlaIrrPheAlaIrrAsn.....PrometAsnG 475  
2001 CTGCG 2050  
475 lProGlyIrrThrGlyGluValIrrProAsnValAlaIrrLeuGlySerProGly 491  
2051 GCGCGCGCGCTATTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2100  
492 LysAlaIrrLeuValAspValSerProAlaIrrThrIrrAspGluArgGlyAl 508  
2101 GCG 2144  
508 rLeuTyrThrArgTyrThrProIrrLeuGlyIrrThrIrrAlaIrrProArgThr 525  
2145 G.....GATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2160  
525 rLeuGlyIrrGlySerGlyIrrGlyIrrGlyIrrGlyIrrGlyIrrGlyIrr 541  
2162 GCG 2211  
542 AlaAspValValValAlaGlyIrrLeuSerGlyIrrLeuGluIrrThrAla 558  
2212 GCG 2260  
558 rGlyValGluIrrAlaGlyIrrLeuSerGlyIrrLeuGluIrrAlaMetAsp 575  
2262 CAAATGATCTTAATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2311  
575 rGlyValGlyIrrGlyIrrThrLeuSerValIrrSerGlyIrrProGlyIrrThr 591  
2312 CTGTAAGCG 2361  
592 lLeuSerAspLeuGlyAlaIrrAlaIrrPheAsnAspSerAsnGluIrrGlyVal 608  
2362 ATCAAGAGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2411  
608 lSerArgLeuPheGluArgPheHisLeuGluIrrGlyIrrLeuGluIrrArgThr 625

```

|||||
2412 ATCGAATTTGTAAGAAATTTCAATTTGGAGGGAGCTCGAGAGGAGCA 2461
625 brgLYASerIleHisGlnAlaGlnAspGlyThrThrThAlaPro 641
|||||
2442 ATGGAATTCATGCTCATCGAGCAAGACGCTACACACACACAGCTCT 2511
642 TGTGASPSerLeuLeuSerGlnGluValAlaSerAlaGlnAlaGlu 658
|||||
2512 ATGAGTATTTCTTCTTACGCAAGAGTTTCAATGCTACTTGGGAACT 2561
458 AluProValTTPSerGluLeuIleGluGlnHisSerLeuGlnAspLeu 675
|||||
2562 GTCGAGGATGAGTCTACGCTGATCGAGAGCAATAGCTTCAGACCTCA 2611
475 YSAIAsrProGlnAlaAlaSerGlnGlnAspSerValSerPheAlaHis 691
|||||
2612 AGGCGAAACCTTAAAGCAACCGGCTCGAAAGTGTACCTTCCGCACTAC 2661
692 TGGAGTGGCTGAGAGAGAGGAGGAGTAAATAGCTGATGATGATGAT 708
2662 TGTGAGAGAGAAATTAATCTGCTGCTGCTTCTTCTGCTGCTGCTAA 2711
708 GThaGAlaLeuLeuGlyValGlnAlaHisGlnIleSerMetLeuPhe 725
|||||
2712 GACAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2761
725 GThaSPSTPLeuGlySerAlaThrGlyLeuSerAsnIlePheSerGAsp 741
|||||
2762 TACGCAATATATGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2811
742 TysGAsPGLyGlyGluThrAlaTGCysLysThrGlyMetIleSerI 758
|||||
2812 AAGAAAGAGAGAGAGAGATATATGATGCAAAAGAGAGAGAGAGAGAT 2861
758 GysHisAlaMetSerLysGlnLeuValTProGlySerValHisLeuAsn 775
2862 TTGGCATGCAATGCAAAAGAACTGTGTCTGAGCTGAGTCACTCAACA 2911
775 hPProValAlaGlnIleGlnGlnSerAlaSerGlyGlySThValArSer 791
2912 GCGCGTGGCGAAATTCAGCAGTGGGATCGGCTGATGAGAGAGAGAG 2961
792 AlaSerGlyAlaValPheArSerLysLysValValIleSerLeuProTh 808
|||||
2962 GAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3011
808 rThLeuGlyThrThrLeuThrPheSerProProLeuProAlaGluLys 825
|||||
3012 AGGCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3061
825 hAlaAlaAlaGlnAsnSerIleLeuGlyTTPTProLysIleValPhe 841
|||||
3062 AAGCAATGGGAGAAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3111
842 ValTTPAspLysProTTPTPArGlnIleHisSerGlyValLeuGln 858
|||||
3112 GTATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3161
858 hAsrSerGlyAspProIleSerIleAlaArAspThrSerIleAspVal 875
|||||
3162 ATGCAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3211
875 sPArGlnITPserIleThrCysPheMetValGlyAspProGlyArGly 891
|||||
3212 ATGCAAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3261
892 TTPSCTGAGGInserLysGlnValArGlnLysSerValTTPAspGln 908
|||||
3262 TGGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3311
908 GArGAlaAlaTyrTrpAsnAlaGlyValAlaGlnValTProGlnProAla 925
|||||

```

```

3312 CGCGCCAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3361
925 sAluGlnIleGluTTPSerLysGlnIleGlyThrThrAlaProPhe 941
|||||
3362 TGTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3411
942 AluValTTPGlnGlnAspLeuIleThrGlySerAlaGlnArGln 958
|||||
3412 GCGCTATGGGCTGAGAGAGATGATCATGACAGTGGCTGAGAGAG 3461
958 TProPheLysSerValHisPheValGlyThrGlnThrSerLeuValTTP 975
|||||
3462 GCGCTTCAAGAGTTCATTTGTTGAGAGAGAGAGAGAGAGAGAGAG 3511
975 YSGTTPGAGTGLuGlyAlaIleArGSerGlyGlnArGlyAlaArGln 991
|||||
3512 AAGGTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3561
992 ValValAlaSerLeuValProAlaAla 1000
3562 GTTGTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

seq_name: us-09-658-835c-25.rni
seq_documentation_block:
  / Sequence: 15, Application: us-08888949
  / Patent No. 6025188
  / GENERAL INFORMATION:
  / APPLICANT: Duvick, Jonathan
  / ARTICANT: Madise, Joyce R.
  / APPLICANT: Wood, Tracy A.
  / TITLE OF INVENTION: FUMONISIN DETOXIFICATION COMPOSITIONS
  / TITLE OF INVENTION: AND METHODS
  / NUMBER OF SEQUENCES: 17
  / CORRESPONDENCE ADDRESS:
  / ADDRESS: PIONEER BI-BRED INTERNATIONAL, INC.
  / STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
  / CITY: Johnston
  / STATE: Iowa
  / COUNTRY: USA
  / ZIP: 50131
  / COMPUTER READABLE FORM:
  / MEDIUM TYPE: Floppy disk
  / COMPUTER: IBM PC compatible
  / OPERATING SYSTEM: PC-DOS/MS-DOS
  / SOFTWARE: Patent In Release #1.0, Version #1.30
  / CURRENT APPLICATION DATA:
  / APPLICATION NUMBER: us-08-888-849
  / FILING DATE: 07-JUL-1997
  / CLASSIFICATION: 435
  / PRIOR APPLICATION DATA:
  / APPLICATION NUMBER: US 08/484,815
  / FILING DATE: 07-JUN-1995
  / PRIOR APPLICATION DATA:
  / APPLICATION NUMBER: US 08/289,595
  / FILING DATE: 12-NOV-1994
  / ATTORNEY/AGENT INFORMATION:
  / NAME: Sweeney, Patricia A.
  / REGISTRATION NUMBER: 32,734
  / ATTORNEY/DOCKET NUMBER: 27283
  / TELECOMMUNICATION INFORMATION:
  / TELEPHONE: (515) 248-4800
  / TELEFAX: (515) 248-4844
  / INFORMATION FOR SEQ ID NO: 15:
  / SEQUENCE CHARACTERISTICS:
  / LENGTH: 1927 base pairs
  / TYPE: nucleic acid
  / STRANDEDNESS: single
  / TOPOLOGY: linear
  / MOLECULE TYPE: DNA (genomic)
  / US-08-888-949-15

```

[illegible]



```

1203 ACCGACACATATCTGAGAGACACATCCGCAATGACCCGACCTTACCA 1252
365 nThrluenglValatyrProtlpeglserProsllytlcllyserProG 382
1254 GATCTCTCTTGGACATATCCCATATGATCCGACGAGATCCGATCGCTC 1402
382 lnsppglnlleAlaAlaAlclclatThrGlnValAArgPheGlnGlyProSer 398
1303 AAGATGAGATTGGGCGCATTCACACGACGCTAGATTCGACGATGCTCTCT 1352
399 AlAlleValAlaGlnPhePheSerAlaArgAsnArgGlyTlclProSerTlpaArgly 415
1353 GCGACGCGGCTGACGAGCTCCGGAATCCGGATACGCTCTTGGCGCTA 1402
415 TTYTYYAAsuAlaThrPheGlnPheSerGlnleuAlaPheProGlySerGlyVal 432
1403 CTACATCAATGCGACCTTGATGATCTGGAGCTTTCGCTGCTGCTGCTGCAAG 1452
432 altyrHsSerSerGlnValGlyGlnValPheGlyThrlyThrlyProValAla 448
1453 TGACACACAGCTCTGAAGTCGAGATCGATGCTTTCGACGATACCTCTGCGCA 1502
449 SerAlaThrAlaLeuGlnAlaAsnThrSerGlySerGlyMetGlnValAlaArg 465
1503 AATGGACCGGCTTCGACGCGCGACGAGAGAGAGAAATGATGACAGAGTCTG 1552
465 pAlaAlaPheAlaAlaAsnProMetAlaAsnGlyProGlyTlPlyGlnValAlp 482
1553 GCGGCTTTCGCAAAACCGCATCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1602
482 rAsnValAlaAlaAlaProGlySerProGlyTlPlyAlaAlaGlnValAlaPhe 498
1603 GCAATGTGCGGCGCTTTCGCTGACGACGACGACGACGACGACGACGACG 1652
499 SerProAlaThrAlaAspGlnArgGlnAlaArgTyrThrAlaArgTyrThr 515
1653 TCTCGACGCAATGACGACGACGACGACGACGACGACGACGACGACGACG 1702
515 rGlnleuGlyThrAlaAlaAlaThrThrThrThrThrThrThrThrThrThr 528
1703 TGAGTGGGACAAATGGCGGCGACGACATTTTGACGACGACGACGATTTGTA 1752
528 SerGlyThrLeuSerLeuGly 535
1753 CCTACATGCGCTTTCGAAACCGACCT 1778
seq_name: us-09-658-835c-25.rn1 us-09-658-758-15
seq_documentation_block:
Sequence 15, Application: us-09-658-758
Patent No. 6238330
GENERAL INFORMATION:
APPLICANT: Davick, Jonathan
APPLICANT: Madock, Tracy R
APPLICANT: Wood, Tracy A.
TITLE OF INVENTION: PROMOTIN DETOXIFICATION COMPOSITIONS
FILE OF INVENTION: AND MANUS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER PROGRAM: P. 31.
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262,758
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/484,815
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: us 08/289,595
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
PREFRENCE/DOCKET NUMBER: 27283
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1937 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-262-758-15

alignment_scores:
Quality: 2000 50 Length: 542
Percent Similarity: 96.863 Percent Identity: 94.096 Gaps: 3
alignment_block:
US-09-658-835c-25 x US-09-262-758-15
Align seq 1/1 LO: US-09-262-758-15 FROM: 1 TO: 1937
1 MetAlaAsnLysHisLeuSerLeuSerLeuPheLeuValleuLeuGlyLe 17
153 ATGCGCGACGATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
203 TCGTTTGGCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252
32 laGlyMetValValGlyThrThrThrThrValProGlyThrThrAlaThr 48
253 CTGGATGTGTGTGCGACGACGACGACGACGACGACGACGACGACGACGAC 302
49 ValSerGlnPheLeuGlnValProPheAlaAlaSerTrpAlaAlaAla 65
203 GTCAGCGACGCTCTGCGCGCTTCTTTCGCGCTTCTGCGCGCTTCTGCG 352
65 apProThrArgProValProThrSerThrProLeuGlnAlaThrAla 82
153 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 402
82 TrpAlaAlaGlyProGlnGlnPheAsnTyrProGlnGlnLeuAlaGly 98
403 ATGCTGCGACGATGCTGCGCGCTTCTTTCGCGCTTCTGCGCGCTTCTG 452
99 lIeThrMetAlaArgPheAsnThrProProSerAlaGlyGlySerG 115
453 ATTACGATGCGCGCTTCAATGATGATGATGATGATGATGATGATGATG 502
115 uAspGlyLeuAsnLeuAlaIleTyrValProGlyThrGlnAsnThrAsn 132
503 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
132 yAlaValMetValTrpTlPlyThrGlyAlaLeuGlnThrGlyGlyTrpAsn 148
553 AAGCGCTCATGTTGGATATATAGGTCGACGCTGGAATATGCTGAGANT 602

```



```

122 ACDCTTCAGCTTCAGACACCTTCCTACGACAAACAGCTTCGCCAGCA 171
421 PheGluAsnLeuGln.....LeuPheProGlySerGluValIleThi 434
172 TCGCCAACTGTCAGCTTGGCTTCGTGATGTCAGCCACTTCACAGCA 221
434 SSerSerGluValIleMetValPheGlyThrTyrProValAlaSerAlaT 451
222 CCGTCACATTCATATCTTCAATCTTCACTATCAAGCCGCAAGACCTCA 271
451 hAlaLeuGlnAlaGlnThrSerLysTyrMetGlnClnAlaIlePAlaAla 467
272 CCGCAATTCGA.....GATGGCTTCAGAGCCATCAC 300
468 PheAlaLysAspProMetAsnGlyProGlyTyrPlyGlnValProAsnVa 484
301 TTTCGCTTCGACAGAGCTC.....CTTCCTGACAGCTGTGGTCCA...G1 341
484 LAlaAlaLeuGlySerProGlyLysAlaIleGlnValAlaSpValSerProA 501
342 GTGGGCTTGTCTTCACCTGAAIACCTCTTTGAGGTTCATGCTCAGC... 387
501 LThrIleAspIleArgCysAlaLeuTyrThrArgTyrTyrThrGlnLeu 517
387 ..... 387
519 GYThrIleAlaIlePheAlaThrPheGlyGlySerGlyGlySerGlySerG1 534
388 ..GGGCTGGTGGC.....GG 401
534 YGlyGlySerLysAspAsnValAlaAspValValValValGlyAlaGlyL 551
402 ATATACATCAATCCATACATATGTCAGCTGTACTGCTGGGCTGGCTGCT 451
551 euSerGlyLeuGlnThrAlaIleValValGlnAlaAlaGlyLeuSerCys 567
452 TGAATGCTTGGATACGGCTACGCTCAAAATCCAGGCGCGCTGTGCTGGC 501
568 LeuValIleGlnAlaIleMetAspAlaValGlyGlyLysThrLeuSerValG1 584
502 CTGCTTCTTGAGAGTATGATGCTGATGAGGAGAGAGCTGTGAGCGTACA 551
584 nSerGlyProGlyArgThrThrIleAsnAspLeuGlyAlaAlaIlePAla 601
552 ATCGGCTGCGCATATGATATATATGATGATGCTGCTGCTGCTGATCA 601
601 sAlaSerAsnGlnSerGlyValSerAlaGlnLeuPheGlnArgPheHisLeu 617
602 ATGATATCAATTAAGGCAAGATATGCAATTTGTTGAAACATTTGATTTC 651
618 GlnGlyGlnLeuPheArgThrThrGlyAsnSerIleHisGlnAlaGlnAs 634
652 GAGGCTGAGGCTTCAGAGCAAGACCTGAAATTCATCATCAAGACACAGA 701
634 P GlyThrThrThrAlaProTyrGlyAspSerLeuLeuSerGlnGlu 651
702 CGCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 751
651 GAlaSerAlaLeuAlaGlnLeuLeuProValIlePheSerGlnLeuIleG1 667
752 TTTAATGTCACCTTCGCGACCTTCCTGCGGATGCTGCTGCTGCTGCTGA 801
668 GlnHisSerLeuGlnAspLeuLysAlaSerProGlnAlaLysAlaLeuAs 684
802 GATATATATATATATATATATATATATATATATATATATATATATAT 851
684 pSerValSerPheAlaHisTyrGlyGlnLysGlnLeuAsnLeuProAlaI 701
852 CAGTGTGACCTTCGCGACCTTCGCGACCTTCGCGACCTTCGCGACCTTC 901
701 AlIleGlnValAlaAsnGlnIleThrArgAlaLeuGlnGlyValGlnAla 717
902 TTCTGCGGCTTATATATATATATATATATATATATATATATATATAT 951

```

```

718 HisClnIleSerMetLeuPheIleThrAspTyrIleLysSerAlaThrG1 734
952 CAGCAATATCAATATATATATATATATATATATATATATATATATAT 1001
734 YLeuSerAsnIlePheSerAspLysLysAspGlyGlyGlyMetArgC 751
1002 TCTATATATATATATATATATATATATATATATATATATATATATAT 1051
751 YsLysThrGlnMetClnSerIleCysHisAlaMetSerLysGlnLeuVal 767
1052 GCAAAACAGGATATGATGATGATGATGATGATGATGATGATGATGATG 1101
768 ProGlySerValHisLeuAsnThrProValAlaGlnIleGlnGlnSerAl 784
1102 CTAAGGCTATATATATATATATATATATATATATATATATATATAT 1151
784 aserGlyCysThrValArgSerAlaSerGlyAlaValPheArgSerLys1 801
1152 ATCTGATATATATATATATATATATATATATATATATATATATATAT 1201
801 YsValValValSerLeuProThrThrLeuTyrProThrLeuThrPheSer 817
1202 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1251
818 ProProLeuProAlaGlnLysGlnAlaLeuAlaGlnAsnSerIleLeuG1 834
1252 CCAATCTCTCCGCGCGGAGAAACAAATATATATATATATATATATAT 1301
834 YTYrTyrSerLysIleValPheValTyrAspLysProTyrPArgGlnG 851
1302 CTATATATATATATATATATATATATATATATATATATATATATAT 1351
851 InGlyPheSerGlyValLeuGlnSerSerCysAspProIleSerPheAla 867
1352 AAGCTCTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1401
868 ArgAspThrSerIleAspValAspArgGlnTyrSerIleThrCysPhe 884
1402 AGATATATATATATATATATATATATATATATATATATATATATAT 1451
884 TValGlyAspProGlyArgLysTyrSerGlnGlnSerLysGlnValArg 901
1452 GCTGCGAGACCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1501
901 InLysSerValTyrAspClnLeuArgAlaAlaTyrGlnAsnAlaGlyAla 917
1502 AAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551
918 GlnValProGlnProAlaAsnValLeuGlnIleGlnTyrSerLysGln1 934
1552 CAATGCTCAAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1601
934 nTYrPheGlnClnAlaProSerAlaValTyrGlyLeuAsnAspLeuIleT 951
1602 GATATATATATATATATATATATATATATATATATATATATATATAT 1651
951 hIleGlnLysAlaLeuArgThrProPheLysSerValHisPheValGly 967
1652 CACTGCGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 1701
968 ThrGlnThrSerLeuValTyrPlyGlyTyrMetGlnIleValIleArgSe 984
1702 ATCTGATATATATATATATATATATATATATATATATATATATATAT 1751
984 rGlyGlnArgGlyAlaAlaGlnValValAlaSerLeuValProAlaAla 1000
1752 GCGGCAACGAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGAGAGG 1800
seq_name: /cgn2_6/psd3ata/4/ina/45c_1001.us 09 45c-168-22
seq_documentation_block:
: Sequence 22: Application US/09352168A

```





```

1252 CCAGCTTTTCCGGGTCACGAAGAACAATTTGCCCAAAATTGTATCATTGC63G 1301
984 GTTGTTGGTAGTACTVAlphaviriflavivirusHepatitisBvirusHepadnavirus 951
1302 CTGCGAIAAGCAAGATACTCTGTGTAAGCATCAATCTGTGAGGCGCAC 1351
851 InglyPheSerGlyValLeuHisSerSerCysAspProIleSerPheAla 867
1352 AAGCGTCTCTGGGCGTCCATCCAATCGACGTGTGAACCCCATCTATTTCC 1401
868 AAASPTSRSTLQASFVALASPFAECLSTSPSESLLETHDCYSPLHE 884
1402 AGSAATACCGACATCGACGTGATGACATGACATGTCATTGCTTTGAT 1451
884 LTVKLTAFPRSGLATILYSTFETRTCSGLISGELYSGLYVALARG 901
1452 GGIQTSGAACCCGACACCGACATCGCTGCCCAACTGCCAACACCTCCACAC 1501
901 INLYSSRVALLTPASPGHLLENARALAIATLCYLCHAMNLAIGLYALA 917
1502 AAAAGICTGTCTTRACCAACTCGGCGACCCCTACAGCAAGCCCCGCCCC 1551
918 GLHVAPRPPHPDVAIAAEVVITWGHIDICLSISTPSERTFGSINCI 934
1552 CAAGTCCCAACCCGACACCGACATCGTGTGAAATCGATGAGTGCGACAGCA 1601
934 nrrrPhgInglYAlaprosErAlaValAlcylrcylchAmnApplmlet 951
1602 GIATLCCAAAGAGATCCGAGCCGCGCTCATAGGCGTGAACACATCTATCA 1651
951 hrLeugLyserAlaleuArqTrPrpheLyseVaHisPhValgily 667
1652 CACGTGATCCGCGCTCGACAAATGCTGTCAAATGATTCATTTCCTTGGA 1701
968 InpGlnThrSerSeValElrThyGScTyrtMetGlnclValAlaArgys 984
1702 ACAGAACATCTTTACTTTTGAAGAGATATATGGAAGGAGGACCATACGATC 1751
984 polyGlnArqGlyAlaAlaCluValAlaValasortLeuValProvalAla 1000
1752 GGATCAACAGAGGTGTTCAGAAAGTGTGGTACAGGTGAGGCTGAGCAGCA 1800

seq_name: Acute Pericarditis in rats by sequencing cDNA-2007
seq_documentation_block:
? Sequence 32, Application US/09352159A
? Patent No. 6211434
? GENERAL INFORMATION:
? APPLICANT: Duvick, Jonathan P.
? APPLICANT: Gilliam, Jacob T.
? TITLE OF INVENTION: Amino Polyol Amine Oxidase
? FILE REFERENCE: 1134
? CURRENT APPLICATION NUMBER: US/09/352,159A
? EARLIER FILING DATE: 1999-07-12
? EARLIER APPLICATION NUMBER: 607092,936
? EARLIER FILING DATE: 1998-07-25
? EARLIER APPLICATION NUMBER: 607135,391
? PARTIAL FILING DATE: 1999-08-21
? NUMBER OF SEQ ID NOS: 40
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 32
? LENGTH: 2490
? TYPE: DNA
? ORGANISM: unknown
? FEATURE:
? OTHER INFORMATION: GST-glyc (AAPAO open reading frame; 2490 nt; GST
? OTHER INFORMATION: and linker; nt 1-687; glye (-) AAPAO; nt 688-2490;
? OTHER INFORMATION: mutation in putative glycosylation sites); all
? OTHER INFORMATION: 1288-1290 (AAAGTCC) and nt 1800-1805 (AAAGTAA).
NAME/Rt.: Glc

```

[illegible]

[illegible]



```
844  GAAATYTTYSerLysIleValIlePheValITPaspLysProITPTrpArg 849
      |||
1945  TGGGTACTAAGCAAGATAGCTTTCGATGGACAGCCGTGGGGCC 2034
      |||
850  GAGAGAGGAGCTGValIleGluIleSerSerCysAspProIleSerPh 866
      |||
2045  GAAATAGGATTTGGAGGAGTCTCCATGTGATGTGATGCCATCTCAT 2084
      |||
866  GATTAATAspThrSerIleAspValAspArgGluITPThrCysP 883
      |||
2085  TGGTAAAGATATCAATGAGAGTCAATGATGATGATGATGATGAT 2134
      |||
884  TCGTAAAGATATCAATGAGAGTCAATGATGATGATGATGATGAT 2184
      |||
2145  TCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2184
      |||
940  ATGAGLysSerValITPAspGluIlePheArgAlaAlaTyrGluAsnAlaG 946
      |||
2185  GAAATAAAGATTTGTGGGAGTCAATCTGGGGTAAAGATGAGATG 2234
      |||
946  YATGAGLysValITPArgIleProAlaAsnValIleGluITPThrLysG 933
      |||
2245  GAGATAAAGATTTGTGGGAGTCAATCTGGGGTAAAGATGAGATG 2284
      |||
943  TGGITTTTThrIleGluValITPAspAlaValITyrGlyLeuAsnAspLeu 949
      |||
2285  AATCATATTTTCAAGAGAGTCTGGAGTGGGTATATGGGTGAGAGCAT 2334
      |||
950  TTTThrLeuGlySerAlaLeuArgThrITPThrLysSerValIleS 966
      |||
2345  ATCAATATGGGTTCAGAGGCTCAAGAAAGGCTCAAGAGGCTCAT 2384
      |||
946  TGTATThrGluIleThrSerLeuValITPLeuGlyTyrMetGluGlyAlaIle 983
      |||
2385  TGGAAAGCAAGCTCTTACTTGGTAAAGAGTATATGAAAGGCTCAT 2434
      |||
984  TGGAGTCTCTAATGCTTAAAGAGTAAAGAGTAAAGAGTAAAGAG 999
      |||
2445  GATTCAGTCAAGAGAGTGTGTGAGAGTGTGTGAGTGTGTGAGT 2484
      |||
1000  Ala 1000
      |||
2485  GCA 2487
```





REMARKS  
COMMENT  
USA  
NIN MOC Di joo + 771 1111 1111  
Contact: moc help desk

USA  
 NIH, MSC B1, Box 361, Bethesda, MD 20892-3610  
 Contact: MGC help desk  
 Email: [cgabst@mail.nih.gov](mailto:cgabst@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgc.bcm.tmc.edu/mgc/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P. P., Garcia, A.M., Liu, X., Huijk, S.W., Hale, E.M., Hahn,  
 Yoon, V.S., Kowis, C.Y., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.

Clone distribution: MCG clone distribution information can be found through the M.A.T. database information at <http://imga.fhn.gov>.  
Series: Iwak Plate-42, Iowa P Column 20  
This clone has the following problems: frame shifted.  
Location: zone 1, east

FEATURES	LOCALITY/CONTINENT
SOURCE	1. 2038

```

/organism="Homo sapiens"
/alt_xref="taxon.6406"
/ncbi_nucleotide="113303"
/issue_type="liver, normal, 5 month old male mouse."
/clone_lib="NCI_CCAP_114"
/alt_host="DH10B"
/feature_vector="FCMV-SPK1E"
BASE COUNT      513 a      535 c      495 g      495 t
ORIGIN

```

```

alignment_scores:
  Quality: 439.00      Length: 519
  Ratio: 1.614         Gaps: 23
  Percent Similarity: 52.108      Percent Identity: 36.220
alignment_block:
  US-09-65b-8350c-25 x RC019426
Align seq 171 to: RC019426      from: 1      to: 2048

```

alignment\_block

US-09-658-835C-25 x R(1)19924

Align seg 1/1 to RC014926 from: 1 to: 2038

```

26 ProthValAlaYlValAspAlaGluMetValValGlyProThrProVal 42
   111 111 111 111 111 111 111 111 111 111 111
119 GCTAAATGATATACGCTGCTGGCTGGCTGTTGATGAGTGGGAGTGGG 168
   111 111 111 111 111 111 111 111 111 111 111
42 1 ProGlyThrThrAlaIleValSerGluLeuGluValProthAla 59
   111 111 111 111 111 111 111 111 111 111 111
169 GAATGATACAGACCCATGATTAAGTGTTCCTGGAGTGGCGATTCCTC 218
   111 111 111 111 111 111 111 111 111 111 111
59 laSerProthr.....AcGPhaAlaProProThrArpProValIro 72
   111 111 111 111 111 111 111 111 111 111 111
219 AACACACAGATGGAGTCTCTGCTCTCACTACGACGCCACACACGCC 268
   111 111 111 111 111 111 111 111 111 111 111
73 IProSerThrProGluIleAlaThrAlaGlyGlyProAlaGlyProGlu 98
   111 111 111 111 111 111 111 111 111 111 111
269 TGGGACGCTGTAGAGATGGCAGATCATCATTCGCCCAATGTTCTTC 318
   111 111 111 111 111 111 111 111 111 111 111
89 .....GluPhaAsnIlyProIleGluIleuAArg 98
   111 111 111 111 111 111 111 111 111 111 111
319 TGTAGAGAAATGATCATACAGCAGAAATTGGCTCATGACAAAGATTA 368
   111 111 111 111 111 111 111 111 111 111 111
98 IuIleThrMetAlaIrrPhaAsnThrProProProSerSerAlaGlyI 114
   111 111 111 111 111 111 111 111 111 111 111
169 ICTTCCCAT.....TCT 382
   111 111 111 111 111 111 111 111 111 111 111
115 GluAspProSerIleuAsnIleuAsnIleIyValProIly.....ThrGluAs 129
   111 111 111 111 111 111 111 111 111 111 111
383 GAGCAGATGCTCATGCTCAAATCTACAGCCCGCATGATGATCATGGAG 432
   111 111 111 111 111 111 111 111 111 111 111
129 nIrrAsnIlyAlaIleMetValIrrIleIyGlyGlyValIleGlyIlyPro 145
   111 111 111 111 111 111 111 111 111 111 111

```

[illegible]





```

244 AlaAlaIleMetGluSerGlyValAlaAspIleAspIleProIleGlyAsp 260
    |||
860 AAGCCGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
260 pIuSerGluProIleAspIleProIleAspIleProIleAspIleProIle 266
    |||
890 .....GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 912
269 .....ValAlaAlaIleAspIleProIleAspIleProIleAspIle 275
    |||
943 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
276 ThrSerIleAspIleProIleAspIleProIleAspIleProIleAspIle 292
    |||
983 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1012
292 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
    |||
1033 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058
409 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 415
    |||
1059 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1108
415 .....SerGluIleAlaAspIleProIleAspIleProIleAspIle 427
    |||
1109 TTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
427 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 444
    |||
1159 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
444 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
    |||
1208 .....GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1216
461 ProAspIleProIleProIleProIleProIleProIleProIleProIle 477
    |||
1217 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1242
477 yIleGly.....SerProIleAspIleProIleAspIleProIleAspIle 489
    |||
1243 ATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
489 ThrGluValIleProIleProIleProIleProIleProIleProIleProIle 495
    |||
1281 AGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1324
406 ArgAspIleProIleProIleProIleProIleProIleProIleProIle 416
    |||
1325 .....GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1351
seq_name: qb_est2:B1332322
seq_definition: block
LOCUS B1332322 805 bp mRNA linear EST 30-JUL-2001
DEFINITION M329850EST.101.2001.210 Mus musculus cDNA clone IMAGE 516456.57.
BRNA sequence.
ACCESSION B1332322
VERSION B1332322.1
KEYWORDS EST.
SOURCE house mouse,
ORGANISM Mus musculus,
Eukaryota; Mammalia; Eumetazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionata; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 805)
AUTHOR National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
1 Dupont, J. (1999)
CONTACT Robert Straussberg, Ph.D.
EMAIL: rstrauss@nigms.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
cDNA library preparation: Life Technologies, Inc.

```

```

FEATURES
    source
        cdna library Arrayed by: The I.M.A.G.E. Consortium (LNL)
        DNA sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LNL at:
        http://image.lnl.gov
        plate: LLM11335 row: 1 column: 01
        high quality sequence start: 2
        high quality sequence stop: 805.
        location/qualifiers
            1..805
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:516456"
            /clone_lib="NCI_CGAP_L19"
            /lab_host="DH10B (TI phage-resistant)"
            /note="Genom. Vector. Vector: pCMV-Sp65; Site: 1; Pol:
            Site_2; Salt: cloned unidirectionally. Primer: 5' to 3'
            Average insert size 1.9 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP library."
            Technology: 216 a 215 a 208 t
        ORIGIN
        EAST COAST 166 a 216 c 215 a 208 t
        ORIGIN
        alignment_scores:
            Quality: 409.00 length: 252
            Ratio: 2.556 gaps: 9
            Percent similarity: 63.492 Percent identity: 46.238
        alignment_block:
        55 59 658 835c 25 x B1332322
        Align seq 1/1 for B1332322 from: 1 to: 805
        12 TGGValLeuLeuGlyGluSerValAspIleProIle.....SerGlu 24
            |||
            |||
            |||
38 GATATATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 87
24 TATATProIleValIleIleAspIleGlyIleValIleGlyIleThrIleT 41
    |||
    |||
    |||
88 ACCAGCTGTGGTGAACACTGTAAAGTAAAGTCTGGAGAAATATGTA 137
41 TrValProGlyThrIleThrIleThrIleProIleProIleGlyValIleProIle 57
    |||
    |||
    |||
136 ATTGAGAAATATGTAACAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 187
58 AlaAlaSerPro ThrArpAlaProProIleArgProVal 71
    |||
    |||
    |||
188 GCGAAGCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 237
71 IPProIleProIleProIleGluAlaThrIleAlaThrIleProIleGlyProIle 88
    |||
    |||
    |||
238 TCTTGAAGATTTGGTGAAGAAATGATTTGATTTGATTTGATTTGATTTGAT 287
68 ThrIleProIleProIleProIleGluAlaThrIleAlaThrIleProIleGlyProIle 103
    |||
    |||
    |||
288 AAGATATTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 337
104 ThrAspIleProIleProIleGluAlaThrIleAlaThrIleProIleGlyProIle 120
    |||
    |||
    |||
338 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 378
120 ThrSerIleProIleProIleProIleProIleProIleProIleProIleProIle 135
    |||
    |||
    |||
379 GATATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 428
429 TATGTCATGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472
152 TGGTATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 168
    |||
    |||
    |||
473 ACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522

```









```

VERSION: B1554561 01:15012286
KEYWORDS: house mouse,
SOURCE: Mus musculus
ORGANISM: Mus musculus
REFERENCE:
  1. (bases 1 to 980)
  NIH Map http://www.ncbi.nlm.nih.gov/Map/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgarps-remail.nih.gov
  Tissue Procurement: Jeffrey E. Green, M.D.
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/Link at:
  http://imgp1111.gsc
  Plater: J1AM1324 Row: K Column: 05
  High quality sequence start: 819.
  Location/Qualifiers
    1..980
      /organism="Mus musculus"
      /strain="V19"
      /db_xref="taxon:10090"
      /clone="IMAGR:112284"
      /lab_hosts="ECHO6 (11 phage-resistant)"
      /note="Organ: liver; Vector: pCMV-Script; Site: 1; Note:
      Average insert size 1.9 kb. Constructed by Life
      Technologies. Note: this is a N1 cDNA library"
FEATURES
  CDS
    220 a 256 c 248 g 246 t
    ORIGIN
      1
      220 a 256 c 248 g 246 t
      1
  alignment_scores:
    Quality: 398.50 Length: 317
    Ratio: 2.202 Gaps: 11
    Percent similarity: 57.098 Percent identity: 37.549
  alignment_block:
    us-09-658-835C-25 x B1527629 ..
  Align seq 1/1 103 B1527629 from: 1 to: 980
  25 ProthylValysIleASPAlcylMcValValcylThrThrThThThVa 42
  121 AAGAGATTTGACAGAGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 170
  71 CTTGGTGGACACACACACAGAGCTCTGGGAGACTATATTCAGTTT 120
  42 ProGlyThrThralaThrValSerGluPheLeuGlyValProPheAlaA 59
  121 AAGAGATTTGACAGAGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 170
  59 IAserPro.....ThrArgPheAlaProProThrArgProValPro 72
  171 AAGAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 220
  73 IAserThrProLeuGluAlaThrAlaThrValProAlaGlyProGlu 89
  241 TCGAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 270
  89 IAserValProGluGluLeuArgGluIleThrMetAlaTrpPheAsot 106
  271 TCGAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 308
  106 IAserProPro.....SerAlaGlyGluSerGlyGlyGlyGlyGly 129
  409 CTAAAGAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 358
  121 AsnIleTrpValProGly.....ThrGluAsnThrAsnIleValAla 135
  1111111111111111111111111111111111111111111111111

```

```

359 AATATTATAGTCCCGGATTTGACAAAGATTTGACAAATTTGACAAAT 408
135 TATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 452
409 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 452
152 euIAserPheAlaSerPheAlaAlaSerPheAlaAlaSerPheAla 168
453 CTATATGACACTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 502
169 TAsnIleTrpValThrValSerGluPheLeuGlyValProPheAla 184
503 ATTCAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 550
184 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201
551 ....CACAGCAGAGAACTGAGCTCTCTCTCTCTCTCTCTCTCTCT 596
597 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 646
218 ThrIlePheGlyGluSerGluPheLeuGlyValProPheAlaAla 234
647 ACCAGCTTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696
234 TAsnIleTrpValThrValSerGluPheLeuGlyValProPheAla 251
697 ATTCAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 746
251 ATTCAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 267
747 TAsnIleTrpValThrValSerGluPheLeuGlyValProPheAla 760
268 ThrValGluAlaLeuAsnGlyThrThrSerIleAsnIleLeuSer 284
761 AACCTGAGCAACAGAGATTTGACAGAGCTCTCTCTCTCTCTCTCT 800
284 eIAserValAlaSerPheAlaAlaSerPheAlaAlaSerPheAla 297
801 .....ATAGCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 846
298 GlnLeuGlyPheGlyPheGlyPheGlyPheGlyPheGlyPheGly 314
837 CAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886
314 TAsnIleTrpValThrValSerGluPheLeuGlyValProPheAla 340
887 CTCTCTGAGAGATTTGACAGAGCTCTCTCTCTCTCTCTCTCTCT 945
seq_name: gb_est2:B1554561
seq_documentation_block:
  LOCUS: B1554561 973 bp mRNA linear EST 05-SEP-2001
  DEFINITION: B032359311 NCL_GCAN_119 Mus musculus cDNA clone IMAGR:1504856 5'
  mRNA sequence.
  ACCESSION: B1554561
  VERSION: B1554561.1 01:15411875
  KEYWORDS: EST.
  SOURCE: house mouse.
  ORGANISM: Mus musculus
  REFERENCE:
    1. (bases 1 to 973)
    NIH Map http://www.ncbi.nlm.nih.gov/Map/
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgarps-remail.nih.gov
    Tissue Procurement: Jeffrey E. Green, M.D.
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
    DNA Sequencing by: Incyte Genomics, Inc.

```

alignment\_block:









Plate: L1AM1113 row: k column: 23  
 High quality sequence strip: 785  
 Location/qualifiers

## FEATURES

Source

/organism "Mus musculus"  
 /strain "FVB/N"  
 /db\_xref "taxon:10090"  
 /clone\_lib "N1\_2AP\_K144"  
 /lab\_host "mlh (T. phaeosporus)"  
 /note "organ: kidney; Vector: pCMV-SpeK; Site 1: Not;  
 Site 2: Salt cloned unidirectionally. Primer: 100 bp dT.  
 Average insert size: 1.75 kb. Constructed by Lile  
 Technologies. Note: This is a N1\_2AP library. 1"

BASE COUNT 207 a 215 c 250 g 224 t  
 origin

## alignment\_scores:

Quality	382.50	Length	264
Ratio	2.27	Gaps	9
Percent Similarity	64.773	Percent Identity	40.909

## alignment\_block:

US-09-658-835C-25 x H100328

Align seq 1/1 to: B1100328 from: 1 to: 878

```

9 SerLeuPhoLeuValIleuLeuGly.....LeuSerAlaSerLeuAlaSo 24
   |||||
24 ACCTGCTGCTGGTGTCTAGCAACCTGTGATTTACCGAAATCCCTC 74
   |||||
23 TGLYALPProThValIlysLeaspAlaGlyMetValValClyThrThr 40
   |||||
74 TTCAACAACCTGGTGGTAACTGATGATGATGATGATGATGATGATGAT 124
   |||||
40 hcthrValProGlyIhrIhrAlaIhrVaSerIhrPhoLeuGlyValPro 56
   |||||
124 TAAAGCTAAAGAGATTTTCACAGCTGCTGGCTCTTCTCTGGAAATCCCT 178
   |||||
57 hcthrAlaIhrPro.....IhrAcPhoAlaIhrProIhrThrThr 70
   |||||
174 TTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 224
   |||||
70 ovalProIhrProIhrProLeuIhrAlaIhrAlaIhrGlyProAlaGly 87
   |||||
224 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274
   |||||
87 ProIhrIhrPhoAsuIhrProIhrIhrIhrIhrIhrIhrIhrIhrIhr 104
   |||||
274 GCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417
   |||||
104 hcthrAsuIhrProIhrProIhrAlaIhrGly.....SerIhrAsp*ysLeuAs 119
   |||||
418 ...AATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
   |||||
119 hcthrAsuIhrValIhrProIhr.....IhrIhrAsuIhrAsuIhrAlaIhr 144
   |||||
364 GCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
   |||||
144 AlMetValIhrIhrProIhrGlyIhrAlaIhrIhrIhrIhrIhrIhrIhr 156
   |||||
415 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 464
   |||||
151 hcthrLeuIhrAspIhrAlaIhrSerPhoAlaIhrAsuIhrAspValIhrAlaIhr 167
   |||||
462 ...AGCTTTCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
   |||||
167 hcthrLeuAsuIhrValIhrAsuIhrLeuLeuLeuLeuLeuLeuLeuLeu 184
   |||||
509 GAGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 557
   |||||
184 GubProIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhr 200

```

```

558 ...GAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 605
   |||||
201 ASPFFVALGILATGASuIhrAlaIhrPhoLeuGlyIhrAspIhrAspIhrVa 217
   |||||
605 CAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
   |||||
217 hcthrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhr 224
   |||||
655 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704
   |||||
234 hcthrMetIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhrIhr 250
   |||||
705 TGTCCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 754
   |||||
250 GVALAlaAsuIhrAspIhrProIhrIhrIhrIhrIhrIhrIhrIhrIhr 263
   |||||
755 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 785

```

...